publisher.agency
Denmark

September, 2026
№ 14

Copenhagen, Denmark
10-11.09.2026

International Scientific Conference

Foundations and Trends in Research

UDC 001.1
P 97

Publisher.agency: Proceedings of the 14th International Scientific Conference «Foundations and Trends in Research» (September 10-11, 2026). Copenhagen, Denmark, 2026. 274p

ISBN 978-3-2247-9642-3
DOI 10.5281/zenodo.22738905

editor@publisher.agency
https://publisher.agency/

Table of Contents

Philological Sciences

Psychological Sciences

Economic Sciences

Technical Sciences

Medical Sciences

Legal Sciences

Political Studies

Pedagogical Sciences

Historical Sciences

Geographic Sciences

Sociological Sciences

Philological Sciences

MEMORY, TIME AND THE SPACE OF HOME IN THE PROSE OF TAUFIK KARIMOV

Mashakova Ainur Kassymzhanovna
PhD Philology, M.O. Auezov Institute of Literature and Art, Leading researcher, Almaty, Republic of Kazakhstan

The creativity of contemporary Kazakhstani writer Taufik Karimov represents a unique artistic space in which personal memory, family history, the fate of generation, and cultural memory of a nation are combined into a unified system of images and motifs. In the represented works of the writer – "Grandfather Makabka," "A Drawing from Childhood", "Hello, My Home...", "Whistle in the Night," "Rain," "On the Nymphenburger Canal", "Once Upon a Festive Evening...," and "Impossible Tenderness and Pain..." – referring to the past becomes not just a means of reproduction of biographical events but also a crucial artistic principle for understanding human existence.

One of the defining characteristics of T. Karimov's prose is the autobiographical basis of the narrative. The author draws on his own childhood, parents, home, familiar places, historical events, and everyday realities of the past. However, the autobiographical nature of his works goes beyond the mere recording of facts. Experiences are transformed into artistic reality, where specific details acquire symbolic meaning, and individual lives become part of a broader historical and cultural experience.

In this regard, the story "Grandfather Makabka" is particularly significant. Already at the beginning of the narrative, childhood memories emerge as a return to a special space, designated by the author as "the world of childhood" [1, p. 13]. The character's awakening from sleep is accompanied not so much by the recording of a specific event but by a sense of returning to a lost time. The summer landscape, the sounds of nature, insects, grandmother, tea, and conversation in the native language gradually shape the atmosphere of a child's world.

Significantly, memory is activated through sensory perception. In T. Karimov's prose, memory is not abstract but rather sensory. Smells, sounds, landscapes, everyday objects, and human speech become unique "keys" to the past. This is why childhood in his works is not perceived as a completed stage of biography. It continues to exist in the adult consciousness and determines their relationship to the present.

In "Grandfather Makabka" the book's motif also plays a significant role. The discovery of a book titled "Grandfather Makabka" becomes an event that connects the realm of childhood with the realm of literature and imagination. Here, literary text becomes part of the protagonist's personal memory. Thus, the author demonstrates that personality formation occurs not only under the influence of family and the surrounding reality, but also through early encounters with literary language.

The motif of memory is further developed in the story "A Drawing from Childhood." Here, the memory revolves around a winter landscape, a courtyard, snow, and a dog. An ordinary scene gradually acquires emotional and symbolic significance. The child perceives the transformation of space under the influence of snow as an almost miraculous event: the autumn courtyard is transformed into a bright and festive space: "An inexplicable feeling of peace, warmth, and, at the same time, delight arose in my soul, delight at the wonderfull spectacle of Nature unfolding before

my very eyes – the change in an instant from one season to another, from autumn to winter..." [2, p. 18].

The episode in which a real painting is connected to an image in a textbook is of a particular significance. The child recognizes in reality what they previously saw in the drawing, and this impression remains in their memory for the rest of their life. Here, one of the characteristics of T. Karimov's artistic thinking is particularly clearly evident: reality is perceived through pre-existing cultural and literary images.

The home becomes the central image of the writer's prose. In "Hello, My Home..." home is presented not simply as a family residence, but as a space for human life and memory. The author connects it with childhood, parents, family history, and a sense of belonging to a particular world. Returning to one's parents' home therefore takes on the character of a return to oneself.

T. Karimov's home is filled with specific details of the past. The summer courtyard, grandmother Abiyem, family belongings, and familiar surroundings form a unique archive of memory. Home is a place where physical space and historical time are inseparable. Several layers of time coexist here: the author's childhood, his parents' story, his family history, and the historical events of the XX century.

Family memory gradually expands beyond the confines of a personal biography. The story mentions Kazan, Gabdulla Tukay, the revolution, the war, and the family's resettlement to Guryev. Thus, the family's history becomes embedded in the history of the country and the region. Historical events are presented through the lives of specific individuals and through the events that remain in the family's memory.

The image of the father occupies a special place in the artistic world of this work. The memory of him is simultaneously linked with love, gratitude, and the pain of loss. The father is presented as a hard worker, possessing "golden hands" capable of performing a variety of tasks and creating the material space for the family's life. It is he who builds the house and lives primarily for others. At the same time, the bright picture of childhood in the work is linked to a tragic event – the father's death. The reference to 1986 disrupts the idyllic finality of the memory and shifts it into a tragic register. Therefore, T. Karimov’s childhood is not exclusively a time of happiness. It is a space in which love, beauty, loss, and a person's first encounter with the irreversibility of time are simultaneously present.

At the end of the work, the meaning of the image of home undergoes a fundamental shift. The real house gradually deteriorates: the courtyard becomes empty, the buildings fall into ruin, and the familiar space of childhood disappears. However, the destruction of physical house does not mean the disappearance of memory. On the contrary, the family continues to exist in the protagonist's inner world, and the father in his memory and consciousness. Therefore, the house in Karimov's prose has a dual nature: it is simultaneously a material space and a spiritual category.

From this perspective, the title "Hello, My Home..." acquires a special meaning. This appeal is directed not only at a specific building, but also at the entire complex of memories associated with it. The house becomes a symbol of the past, to which people return long after it has disappeared. Its existence continues thanks to memory.

Landscape plays an equally important role in T. Karimov's artistic system. Nature in his works is not a neutral backdrop for events. It is directly linked to the protagonist's emotional state and mechanism of memory. This is particularly evident in the story "Whistle in the Night".

The action of "Whistle in the Night" takes place in the village of Kulsary in the Guryev region in the late 1950s and early 1960s. The semi-desert landscape, poverty, and peculiarities of everyday Soviet life create a specific historical and everyday setting for the narrative. A young teacher, who finds himself in the care of a Kazakh family, gradually integrates into local life, adopts its customs, and finds himself integrated into a multinational cultural environment.

The motif of a mysterious whistle in the night initially creates a sense of mystical menace. The recurring sound disturbs the tranquility of the night and compels the protagonist to search for its source. However, the solution turns out to be quite real: the whistle is produced by a wooden shepherd's whistle, the sound of which is generated by the wind.

This technique has significant artistic significance. The writer creates a sense of mystery, but then returns the reader to material reality. Thus, the everyday episode acquires a philosophical subtext: the most usual thing can become mysterious if a person perceives it through fear, memory, and imagination.

The author's final commentary is also important, recalling teachers and students of different nationalities, as well as the subsequent dispersal of this world after independence: "Among the school teachers were Russians, a Ukrainian, a Latvian, a Jew, a Kazakh, a Tatar, and an Armenian. At the same time, in our class alone, children of nine nationalities studied together – Kazakhs, Russians, Bulgarians, Germans, Tatars, a Ukrainian, an Ingush, a Greek, and an Armenian. And on the school football team of 11 players, which became the best in the city in those years, the same Kazakhs, Russians, a German, a Chechen, and a Tatar competed equally for a single victory... Although, we did not pay much attention to nationality then. Only now, recalling those years, you suddenly see such a diversity of faces and surnames" [3, p. 23]. Thus, a private history turns into a reflection on the past of a multinational society. Memory becomes a way of preserving the lost human space.

A special type of historical memory is presented in the work "Once Upon a Festive Evening....". The events take place in November 1967, during the celebration of the 50th anniversary of the October Revolution. The official symbolism of the holiday clashes with the real-life conversations of adults about food products, prices, a new store, and everyday problems. This situation allows the author to demonstrate the discrepancy between official historical discourse and everyday human consciousness. History in the work exists not in the form of slogans, but through everyday conversations, expectations, and doubts.

The multilingual environment is particularly significant: Russian families, a Kazakh woman, conversations in Russian, Kazakh, and Tatar. The characters understand each other despite the linguistic differences. This forms a model of cultural space, important to T. Karimov, in which various national traditions coexist. In this regard, his prose has not only autobiographical but also cultural and historical significance.

The juxtaposition between the official worldview and a person's inner sense is particularly evident in the image of distant lights. They are perceived as a sign of a different, more prosperous life, while adults are skeptical of official promises. This seemingly small episode becomes a means of conveying the atmosphere of the era through a child's perception.

Several works were written in Munich, where the writer lived in 1993-94. Impressions of this German city are reflected in the stories "...And All the Colors of the Rainbow" and "On the Nymphenburger Canal". The novella "Rain" is dedicated to Munich. The narrator dreams of a quiet, sunny city, but the reality is the opposite: endless rain transforms the familiar space into a zone of anxiety and destruction. The names of specific city landmarks – Sendlinger Tor, Sonnenstrasse, Stachus, Marienplatz – create a recognizable geography. However, gradually, the familiar space becomes alien and threatening as the flooding intensifies. From a literary perspective, "Rain" can be seen as a kind of dystopian miniature. The city, a symbol of modern civilization, proves to be not eternal, but vulnerable.

In the story "On the Nymphenburger Canal" the space of Munich takes on a completely different meaning. The frozen canal, people enjoying winter, and white swans create a harmonious picture of the present. However, this present is suddenly shattered by a memory of the distant Urals, where the character, at the age of twelve, found himself alone on an old skating rink in the cold and darkness. Thus, one space evokes another. Munich and the Urals are

connected in the character's mind through memory. Geographical distance is no longer an obstacle: the past emerges within the present. The final appearance of a pair of white swans reinforces the motif of beauty and simultaneously leaves a sense of ambiguity.

This organization of time is characteristic of all of Karimov's prose. The past is not solely "behind" the protagonist. It constantly returns, entering into dialogue with the present. As a result, the temporal structure of the works becomes nonlinear: the present triggers memory, memory returns to childhood, and childhood is reinterpreted from the perspective of an adult.

This problem is expressed most philosophically in the story "Impossible Tenderness and Pain..." dedicated to his parents, Akram and Rauza Karimovs. Here, memory is completely free from a specific plot event and transformed into a reflection on the nature of parental love. Years later, the person once again experiences the tenderness and pain associated with childhood.

The author associates this feeling primarily with parental love, which is absorbed by the child and then passed on to the next generation. In the finale, parents acquire an almost universal significance: they are presented as the earthly creators of the human world. Memory of them, love for them, can make the world purer, brighter, kinder, and gentler.

Thus, in T. Karimov's prose, memory serves not only the function of preserving the past but also a moral function. To remember means to preserve the connection between generations, to restore meaning to vanished people and events, and to resist the destruction of time.

The uniqueness of T. Karimov's artistic poetics is determined by the combination of several principles. His works feature the documentary concreteness of autobiographical narrative, lyrical emotion, expressive landscapes, symbolic subject matter, and philosophical generalizations of endings. At the same time, the author avoids excessive declarative nature: important historical and moral meanings are revealed through particular situations.

The detail plays a special role. A book, a drawing, a house, a yard, snow, a whistle, night lights, a canal, swans – these are not simply elements of the material world. They become signs of memory. Through them, the past can exist in the present. In T. Karimov's prose, an object often becomes a unique material carrier of human time.

No less significant is the motif of the road and movement between spaces. The characters find themselves connected to different geographic locations – the Urals, Guryev, Kulsary, Kazan, Munich. But these spaces are not mechanically contrasted. They are united within a single internal space of memory. Therefore, the geography of T. Karimov's prose is simultaneously concrete and symbolic.

Another important feature is the multinational cultural context. Russian, Tatar, and Kazakh linguistic and cultural environments coexist naturally in his memoirs, as part of everyday life. This is particularly noticeable in episodes related to family interactions and Kulsary. This depiction of multilingualism allows us to consider Karimov's prose as an artistic testimony to a specific cultural space in Western Kazakhstan.

The image of home is especially important. A physical home may collapse, but the spiritual space of home, created by love, family memory, and human relationships, endures. This is why the final tones of T. Karimov's works, despite the presence of pain and loss, are not confine to tragic pessimism.

Ultimately, all the main motifs – home, childhood, parents, nature, historical memory, the meeting of past and present – are united in the works of T. Karimov by a single central issue: the issue of human memory in the face of time. Time destroys homes, it changes cities, carries people away, and erases the everyday details of the past. However, artistic expression is capable of resisting this disappearance. T. Karimov's prose can be read as a unique literature of memory. Its subject is not only the events of the past, but also the process of their return to consciousness. The writer demonstrates that memory is not a passive repository of the facts. It constantly reinterprets the live experience, connecting personal and historical, the present and the past, the

concrete and the symbolic. In his artistic world, memory becomes a form of moral resistance to oblivion. By remembering parents, childhood, home, and the people and places of the past, a person preserves their own identity and simultaneously restores the past to its human significance. This is one of the main artistic and philosophical values of T. Karimov's prose.

References:

1. Karimov T.A. Grandfather Makabka // Karimov T.A. And all the colors of the rainbow. – Izhevsk: Kama Publishing House, 2000. – pp. 13-16.

2. Karimov T.A. Drawing from Childhood // Karimov T.A. And all the colors of the rainbow. – Izhevsk: Kama Publishing House, 2000. – pp. 17-19.

2. Karimov T.A. Whistle in the Night // Karimov T.A. And all the colors of the rainbow. – Izhevsk: Kama Publishing House, 2000. – pp. 20-23.

CONNECTED SPEECH AS A FACTOR OF LISTENING DIFFICULTIES IN DIGITAL DISCOURSE: A PILOT STUDY ON KAZAKHSTANI EFL LEARNERS

Alieva Dilnaz Maratkyzy
Fourth-year student, Department of Foreign Philology, Faculty of Philology, Eurasian National University, Astana

Abstract

The present study examines how Kazakhstani EFL learners with self-reported B1–C1 English proficiency perceive and decode connected speech in authentic digital discourse. Existing research relies mostly on scripted audio, therefore this paper uses unedited TikTok videos instead. Such content reflects real-world speech conditions more accurately. Sixteen third-year students of the "Foreign Philology" faculty transcribed six audio phrases which included different connected speech processes — elision, yod-coalescence, assimilation and lexical reduction. Additionally, each phrase was analyzed acoustically in Praat in order to check whether the process actually occurred in the signal, rather than being assumed from the transcript. Results showed that even advanced learners systematically had problems decoding a few phrases while still remaining confident in wrong answers. The most striking finding was that when several phonological processes overlapped in one phrase, the difficulty did not simply add up. Instead, it produced a completely different, non-existent word in the listener's perception — and this was visible on the spectrogram itself. Perceptual failure turned out to depend on how deeply speech was phonologically reduced and on how familiar the listener was with the underlying word. By contrast, language proficiency level itself poorly explained these results. The study claims that university language programs need to bring authentic digital discourse into the classroom, in order to close the gap between what students are taught and what they actually hear in real life.

Keywords: connected speech, phonological decoding, EFL listening, acoustic verification, authentic digital discourse

Introduction

Recent shifts in English language pedagogy have moved from mimicking native-like accent as a goal to prioritizing functional intelligibility [5]. This shift demands students to recognize unedited fast connected speech patterns rather than articulated scripted textbook forms. This is a specific challenge for Kazakhstani students who learn English as a foreign language: rhythmic and reduction patterns of their native languages do not map onto English connected speech, therefore students can face phonological decoding failure even when their vocabulary is sufficient.

Existing research still relies on scripted audio materials although scholars call for a greater focus on real-world intelligibility.

For this reason, the goal of this paper is to investigate how Kazakhstani EFL learners perceive and decode connected speech in authentic digital discourse materials.

1. Does the academic level influence the ability to decode phonetic boundaries in unedited digital discourse?

2. What connected speech processes are the most difficult to decode for Kazakhstani students specifically?

3. Do the connected speech processes that create the main perceptual difficulties appear as a measurable acoustic reduction, or does it depend mainly on whether the listener is familiar with the particular word?

Theoretical Foundation

The study of English phonetics has experienced a significant paradigm shift: from trying to copy a flawless traditional accent to global intelligibility. David Crystal claims that the main goal of a modern learner is not necessarily to replicate Received Pronunciation, but to navigate effectively in global communication [5, p. 173]. To explain this transition, R. Cauldwell introduces a difference between “The Green House” (the controlled environment of the classroom) and “The Jungle” (unpredictable reality of natural language) [4]. Connected speech is a basic mechanism of this reality: it emerges due to the fact that native speakers prioritize rhythmic organization, not clear articulation of individual words.

This rhythmic organization is linked to the term isochrony. English is usually described as a stress-timed language following a repeated “beat” between stressed syllables. Kazakh is an agglutinative Turkic language with well-developed vowel harmony, and is usually classified as syllable-timed, with comparatively little reduction of unstressed vowels. Russian, which is considered stress-timed as well [12; 1], reduces unstressed vowels according to its own specific model: for example, unstressed /o/ and /e/ shift toward /a/ and /i/. More recent studies question the strict binary classification of rhythmic classes [13; 6], however, the practical implication for learners remains valid: for native speakers of both Kazakh and Russian the specific way English “swallows” words to preserve its rhythm remains unfamiliar. This holds even for Russian speakers, whose first language reduces vowels in its own way.

Because of this difference in rhythm, switching to real-world English is a major challenge for Kazakhstani learners. A. Aiganatova notes that English rhythm and stress models present a considerable difficulty because native phonetic patterns conflict with the stress-timed nature of English [2, p. 389]. Words are perceived as fortresses with protected boundaries, but in English these “walls” are demolished in order to save time. N. Z. Teshaboyeva and M. Khaidarova point out that this specific process is a key component of spoken fluency [14, p. 178]. Classroom observations confirm this. Trying to find a distinct boundary in such phrases as “stop it”, listeners frequently hear the sound “pit”. This inevitably leads to comprehension breakdown. Even though students make noticeable progress in learning English, connected speech processes remain a persistent obstacle. As a result, high competence itself cannot guarantee proficiency with these phonetic patterns [2, p. 389].

Building on these theories, it is possible to suggest that phonological decoding does not exist without a stable “linguistic anchor” in the learner’s mental lexicon. Personal observations suggest that even when a student has learned the rules of assimilation (for example, “goodboy” turns into /gʊbːaɪ/), it is often impossible to physically “hear” a word without high lexical familiarity with it. The word seems to become acoustically invisible in this case. The learner cannot recognize a distorted sound when no stored visual form of the word exists. However, the acoustic blindness phenomenon appears when the lexical unit is known too. According to G. Lindsey, contemporary phonological processes such as glottalization and yod-coalescence are standard characteristics of modern British English. Yet they appear to create a significant perceptual load for students whose expectations are based on old spelling patterns [9, pp. 9–10, 59, 91]. In this case the brain filters the signal as “white noise”, because it cannot find an ideal lexical match.

Recent studies suggest that listening difficulties emerge from phonological decoding failure, not from a lack of vocabulary [3, p. 443]. When a student faces assimilation, their inability to recover the sound to its written form leads to speech segmentation difficulties. This may explain why many students find it hard to watch movies in the original without subtitles. The speed of the

speech stream seems to erode word boundaries, and because of such phonetic pressure the brain fails to restore "weak" words from memory.

Despite all existing studies, there is still a significant gap in identifying how students handle unscripted, natural audio. N. Momen and Z. Pilus criticized using scripted academic recordings, arguing that they do not prepare students for real-life discourse [10, p. 389]. In contrast to them, this paper fills this gap by using unedited short TikTok recordings. Every phrase has been checked acoustically with spectrogram analysis via Praat — to see whether phonological reduction has actually occurred in the signal. The study asks whether university programs prepare students to recognize connected speech processes in a fraction of a second, or whether students continue to look for "fortress boundaries" that do not exist. If the acoustic verification confirms that a phonological transformation has really taken place, and a student knowing the word in its written form still could not recognize it, this serves as empirical evidence of acoustic blindness.

Methods and Materials

In this paper a mixed-methods approach is used, combining quantitative analysis of perceptual accuracy with qualitative acoustic analysis. The goal is to study how third-year students decode authentic connected speech and to verify acoustically which phonological processes are present in each stimulus.

Participants. Third-year students of the "Foreign Philology" program with self-reported language proficiency (B1–C1) have taken part. The group was chosen deliberately: if even advanced learners struggle with phonological decoding, this is likely to indicate a gap in educational programs rather than a mere lack of vocabulary. Participation was voluntary; all the students were informed of the purpose of this study first.

Stimuli and materials. The six audio materials were taken from TikTok. The platform was chosen because its content contains unedited and unscripted recordings. This makes such stimuli closer to natural everyday speech than textbook recordings. The following phrases were selected in order to cover different connected speech patterns.

Phrase 1: ("It's Tuesday, isn't it?") — this one demonstrates elision and linking: 'isn't it' reduces to /izntit/.

Phrase 2: ("The beauty of a king is that it is for everyone, cause imagine the alternative") — this phrase shows how function words reduce to schwa until word boundaries effectively disappear.

Phrase 3: ("I brought you my old maternity clothes") — the third expression is aimed at yod-coalescence, specifically at the transition of "brought you" to /brɒtʃuː/.

Phrase 4: ("She's been talking about you all summer") — this one contains weakened stop realization at the intersection of 'about you' and vowel linking in 'you all'.

Phrase 5: ("I'm gonna be fine") — this phrase presents a classic example of reduction through contraction, when 'going to' turns into 'gonna'.

Phrase 6: ("Imagine if I have to cancel shows cause I'm having a baby") — overlaps a few processes simultaneously, including assimilation, elision and lexical reduction, whereas 'have to' becomes /hæftə/.

Procedure. Each phrase was played two times with a five-second pause. After this, participants had only one minute to write down their answers. This time limit was adjusted following a pilot study in order to better reflect natural processing.

Acoustic verification and data analysis. Each of the six phrases was analyzed acoustically in Praat. Waveforms and spectrograms were segmented into TextGrids to verify whether the claimed process — elision, yod-coalescence, assimilation, or lexical reduction — was actually present in the signal: whether a word boundary was truly absent, whether a vowel was in fact reduced to schwa. Word frequency was evaluated according to the SUBTLEX-UK database [15] and converted to the Zipf scale. The student responses showed two patterns of answers. The first one is

perceptual misidentification: an incorrect transcription with a confidence rating of 5/5. This indicates that the brain projected the input signal onto a closer form of a word present in one's vocabulary. The second one is uncertain recognition: a correct transcription with a low confidence level. It suggests weak decoding automation. The final question on the most difficult word or fragment allows for identifying which connected speech patterns caused the greatest challenges for Kazakhstani students in particular.

Google Forms structure. The questionnaire was divided into two parts. In the first part, each audio phrase was given its own section. English language level was evaluated at the beginning of the survey. Next, each section had two questions where the students were meant to record how they heard each phrase and to note their confidence level. In the second part, all of the six expressions were shown to the students in their written form, so they could check their understanding. Based on each of these expressions, the participants were asked to indicate a word or a part which was the most challenging in terms of comprehension.

Results and Discussion

Sixteen third-year students of the "Foreign Philology" program with self-reported English level participated in this experiment: B1 — 37.5%, B2 — 50%, C1 — 12.5%. All the participants transcribed six authentic audio fragments from TikTok containing different connected speech processes. The results have shown significant differences in perceptual accuracy and two consistent deviations: perceptual substitution and uncertain recognition.

Phrase 1: "It's Tuesday, isn't it?"

Praat analysis detected no amplitude drop and no boundary between "isn't" and "it": the two words were pronounced as a single continuous sound. The second discovery was evident at the beginning of "Tuesday". Instead of a pure /t/, the spectrogram showed the noisy release characteristic of an affricate — the acoustic picture of yod-coalescence. This particular finding appears to explain the most revealing group of answers. A few listeners wrote down "choose day": the affricate that emerged from coalescence was heard as the beginning of the word "choose". The error is unlikely to be random, as it follows directly from what is present in the signal (see Figure 1).

Phrase 2: "The beauty of a king is that it is for everyone, cause imagine the alternative"

According to the acoustic analysis, the reduction was focused on function words rather than evenly distributed throughout the phrase. Four transitions had no visible boundary: "of" and "a" (schwa reduction), "that it is" (elision plus reduction), "for" and "everyone" (schwa reduction), "cause" and "imagine" (boundary loss). The boundary before the content word "king" remained distinct.

The answers mainly align with this acoustic map. The students lost the middle of the phrase, where the function words were, and recovered its final content word "alternative". Where the signal preserved a distinct boundary, perception appears to have worked; where the boundaries vanished, it failed (see Figure 2).

Phrase 3: "I brought you my old maternity clothes"

Praat analysis confirmed complete yod-coalescence between "brought" and "you". The /t/ was released as a short burst of high-frequency noise (approximately 60 ms, 3200–4200 Hz) — the acoustic signature of the affricate /tʃ/ — and there was no boundary between the two words, therefore the /t/ in fact became the beginning of "you". The next segment, "you my old", was also pronounced as a single continuous sound, while the boundary before "clothes" remained clear.

However, the answers revealed a different kind of failure here. The word "maternity" — Zipf 3.67, the lowest-frequency word in the whole stimuli set — was fully present in the signal. But only two out of sixteen students could restore it. The others replaced it with more familiar words similar in form: "training class", "raincoats". The word was in the signal, but the listeners lacked a

stored form to match it against. This is lexical unavailability, and it can serve as strong evidence that perception depends on the listener's stored vocabulary as much as on the signal itself.

Phrase 4: "She's been talking about you all summer"

The acoustic analysis identified two things here. At the intersection of "about you" the plosive was weakened: energy dropped by approximately 75%, but complete frication of a merged affricate did not arise. Between "you" and "all" the reduction was full — two vowels were pronounced without a boundary between them, as a single linked sound.

This linked "you all" happened to be the main failure point. None of the students divided it into two words. The most frequent answer shows what they did instead ("she's been talking about last summer", 5 out of 16). They retained the acoustically distinct parts of the phrase ("talking about … summer") and filled the fused segment with a familiar word appropriate to the context.

Confidence data show that this phrase was perceived differently than the first one. There, the mode was 5/5, because the listeners heard the incorrect version and were confident in their wrong answers. Here the mode is 1/5 (43.8%). It means the participants decoded the phrase only partially and did understand that some part of it was lost. As a result, a fully merged boundary produced confident mistakes, while a partially reduced one left the listeners unsure. The reduction depth determined both what the students heard and to what extent they believed it.

Phrase 5: "I'm gonna be fine"

The phrase was pronounced as a single continuous interval. The analysis showed no amplitude drop at the level of a boundary, and the whole sequence "I'm gonna be" took about 260 ms. "Gonna" was present in the signal only in its reduced form, without a trace of the full "going to". At 364 WPM, this was the fastest stimulus in the set — almost three times faster than the norm for instructional audio, 100–130 WPM [7].

The results were the worst in the set. There was no correct transcription, only a single answer that contained something relevant ("Oh it's fine"), and responses such as "follow this way" and "anxiety", which did not align with the phrase at all. This failure may be informative precisely because the phrase includes only high-frequency words. "Gonna" (Zipf 5.24) and "fine" (5.44) are among the most frequent items in this stimuli list. The students were familiar with all the words in their written form, so the results cannot be explained by vocabulary. The students lacked the reduced spoken form of the word itself as a recognizable unit. This suggests that word knowledge and its realization in connected speech may be stored separately.

Phrase 6: "Imagine if I have to cancel shows cause I'm having a baby"

The spectrogram showed continuous formants during two extended stretches: "if I have to" and "cause I'm having" were pronounced without measurable boundaries. A few reduction processes overlapped within a single utterance, making it the densest stimulus in the set.

The answers reflect this density. 37.5% of respondents left this item blank — the highest such rate in the set. Most of the remaining answers included only the final content word "baby". This pattern repeats the situation seen in Phrase 2, but more intensely: when the middle of an utterance collapses into a single continuous reduced segment, listeners fall back on the final content word that was articulated distinctly (see Figure 3).

However, one factor here needs to be named. The speaker's British Cockney accent may have added further difficulty, so accent and reduction cannot be fully separated for this phrase. This overlap is discussed in the Limitations section.

The Spearman correlation between language proficiency level and transcription accuracy was weak and statistically insignificant ($\rho = 0.190$; $p = 0.481$; $n = 16$). Language proficiency level was a poor predictor of connected speech perception in this sample.

Across the six phrases, errors fall into four types. Perceptual substitution: the listener hears a different real word built from the same acoustic material ("choose day", "last summer").

Phantom insertion: a word with no grounding in the target occurs instead ("anxiety", "follow this way"). Lexical unavailability: the word is present in the signal but absent from the listener's stored vocabulary ("maternity", replaced with "training class" and "raincoats"). Full collapse: no usable transcription at all (blank responses for Phrases 5 and 6).

Notably, all four error types appear within the homogeneous B2 group, implying that none of these categories is reducible to proficiency level. The considerable variation in responses within a single proficiency band is consistent with models of individual differences in L2 phonetic acquisition. One such model is the revised Speech Learning Model (SLM-r), proposed by Flege and Bohn [18]. Factors behind this variation (for example, phonological working memory or the amount of listening experience) were not measured in this specific design. They deserve further study.

Confidence ratings split into two failure states. In Phrase 1, the mode was 5/5: half of the respondents were certain of their answers, including those who gave "choose day" as correct. Across the whole dataset, approximately every fifth incorrect transcription (7 out of 34, 20.6%) involved the maximum confidence rating — the students often did not realize their perceptual failures. In Phrases 4–6, by contrast, the mode dropped to 1/5: respondents decoded the phrases only partially and knew it. Reduction depth thus produced two distinct perceptual states: complete boundary loss led to confident, incorrect perception, while partial reduction led to acknowledged uncertainty.

Speech rate itself did not determine difficulty. Phrase 2, the second-fastest stimulus in the set (275 WPM), had the lowest failure rate; Phrase 3, pronounced at a moderate 232 WPM, was failed by 81.2% of listeners. The difference between these two stimuli lies in composition, not speech rate: Phrase 3 contained the lowest-frequency word in the set, while Phrase 2 allowed listeners to rely on distinct, bounded content words.

Word frequency interacted with phonological reduction. The lowest-frequency word in the list, "maternity" (Zipf 3.67), was failed even though the signal retained it. Phrase 5, consisting entirely of high-frequency words, was unsuccessful because the reduced form of those words was unfamiliar. Successful perception appears to have required two things at once: a preserved lexical form of the word and familiarity with its reduced spoken realization. This is consistent with classic findings on lexical influence in speech perception. The Ganong effect shows that an ambiguous sound tends to be heard as the one which forms a real word [8]. Warren's phonemic restoration effect works the same way: listeners "hear" a sound that is actually missing from the signal [17]. More recent work points to the same pattern: Mulder and colleagues found that reduced word forms are processed more slowly and less accurately even by native speakers, and especially by L2 listeners [11].

These results are also relevant to listening instruction methodology. Teaching metacognitive strategies [16] enhances listening comprehension through planning and monitoring and retains its value. Nevertheless, the empirical data demonstrate failures that no strategy can correct: a listener cannot track a word that was never decoded in the first place. Bottom-up training in decoding reduced forms, grounded in authentic materials, appears to be a necessary complement [7].

Limitations

Four limitations should be named. First, the sample size was relatively small (n = 16) and was drawn from a single institution, so the correlational results should be regarded as preliminary. Second, each phrase was produced by a different speaker, making it impossible to fully isolate accent from reduction; this is particularly relevant for Phrase 6, recorded in a British Cockney accent. Third, the recording quality of the fifth phrase was lower than that of the other stimuli (RMS 0.19; SNR 25.6 dB vs. 31–33 dB), which could have added to its failure rate alongside

phonological reduction. Fourth, language proficiency level was self-reported, with no standardized test.

Conclusion

The aim of this study was to investigate why Kazakhstani EFL learners experience connected speech challenges in authentic digital discourse. Acoustic analysis confirmed that the problems are rooted in the signal itself: elision, yod-coalescence, vowel linking, and function-word reduction were in fact present on the spectrograms, and the respondents' mistakes followed these processes directly. Simultaneously, the response data reflect that the signal is only half of the problem. Language proficiency level was a poor predictor of accuracy ($\rho = 0.190$; $p = 0.481$). The lowest-frequency word in the set was restored by only two listeners, even though it was fully present in the signal. Conversely, the phrase made up entirely of high-frequency words yielded no correct transcription, because its reduced form was unfamiliar. As a result, connected speech perception appears to depend on the interaction of two factors: the depth of reduction in the signal and the listener's knowledge of the word itself as well as its reduced form. For teaching, this means that instruction requires systematic work with authentic digital discourse materials, alongside explicit bottom-up decoding training [7] together with guidance in metacognitive strategies. The same mechanism — a gap between knowledge of a form and access to it in real time — extends beyond the foreign-language classroom. Studying this in speech production is a natural next step for this research.

Table 1. Acoustic Properties of the Stimuli

	Target Phrase	Main Connected Speech Processes	Words	Dur. (s)	WPM	Speaker
1	It's Tuesday, isn't it?	Elision, linking, yod-coalescence	4	1.97	122	F / GB (regional)
2	The beauty of a king is that it is for everyone, cause imagine the alternative	Schwa reduction, elision, function-word boundary loss	15	3.27	275	M / GB
3	I brought you my old maternity clothes	Yod-coalescence, vowel linking	7	1.81	232	F / GA
4	She's been talking about you all summer	Weakened stop realization, vowel linking	7	1.77	237	M / Standard Southern British
5	I'm gonna be fine	Lexical reduction, phrase-final decay	4	0.66	364	F / GA
6	Imagine if I have to cancel shows cause I'm having a baby	H-dropping, function-word elision	12	2.12	339	F / GB (Cockney)

Note: Duration measured via Praat acoustic segmentation (Boersma & Weenink); WPM = (words / duration in seconds) × 60. Standard L2 classroom speech rate: 100–130 WPM [7].

Table 2. Student Transcription Accuracy and Confidence Ratings

	Target Phrase	Exact n/16 (%)	Partial n/16 (%)	Wrong n/16 (%)	Blank n/16 (%)	Mean Conf.
1	It's Tuesday, isn't it?	9 (56.2%)	3 (18.8%)	4 (25.0%)	0 (0.0%)	4.06
2	The beauty of a king is that it is for everyone, cause imagine the alternative	0 (0.0%)	15 (93.8%)	0 (0.0%)	1 (6.2%)	2.13
3	I brought you my old maternity clothes	2 (12.5%)	1 (6.2%)	11 (68.8%)	2 (12.5%)	2.38
4	She's been talking about you all summer	0 (0.0%)	10 (62.5%)	6 (37.5%)	0 (0.0%)	2.25
5	I'm gonna be fine	0 (0.0%)	1 (6.2%)	13 (81.2%)	2 (12.5%)	2.38
6	Imagine if I have to cancel shows cause I'm having a baby	0 (0.0%)	10 (62.5%)	0 (0.0%)	6 (37.5%)	1.50

Note: Exact = fully correct transcription of the entire target phrase. Partial = at least one key content word decoded correctly, but the transcription is not fully accurate. Wrong = no target content word recovered. Blank = no response given. Confidence was rated on a 1–5 scale.

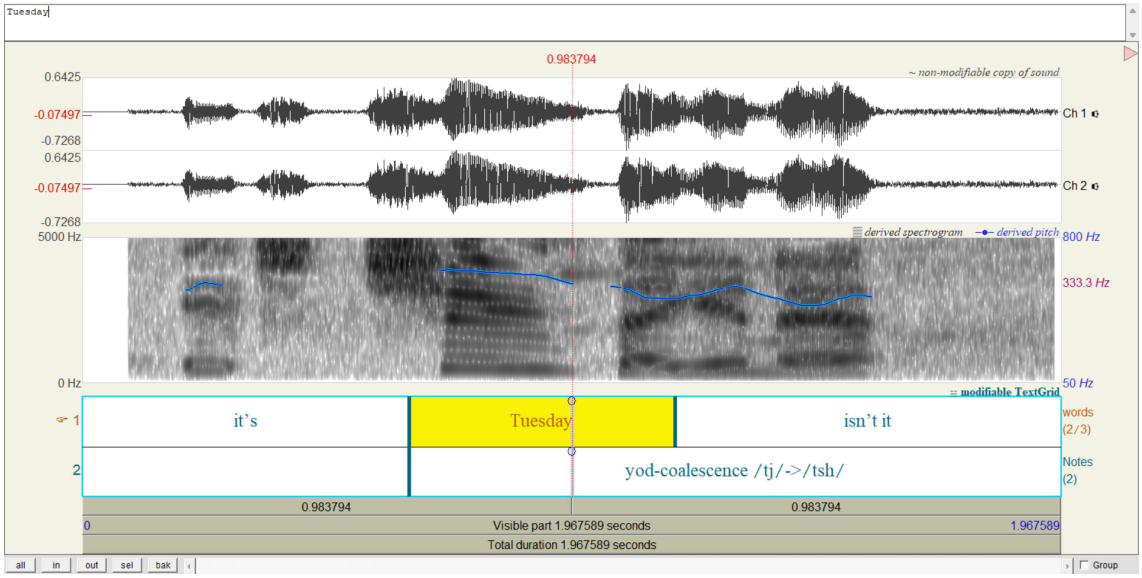


Figure 1. Spectrogram of Phrase 1 ("It's Tuesday, isn't it?"): the absence of a boundary between "isn't" and "it", and the noisy onset of "Tuesday" consistent with yod-coalescence.

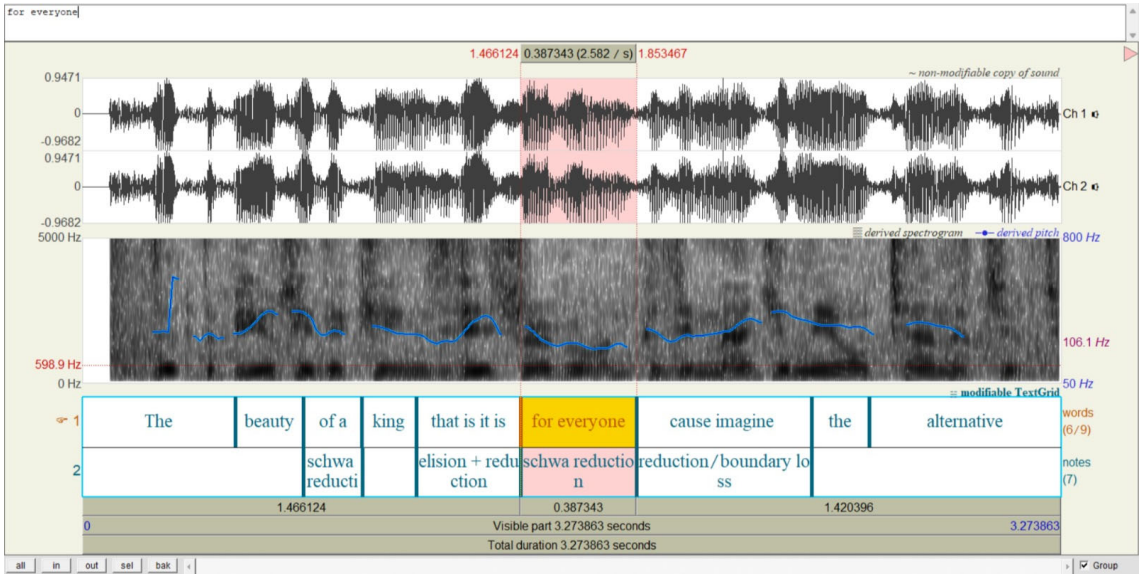


Figure 2. Spectrogram of Phrase 2: four points of boundary loss between function words ("of a", "that it is", "for everyone", "cause imagine"), compared with the preserved boundary before the content word "king".

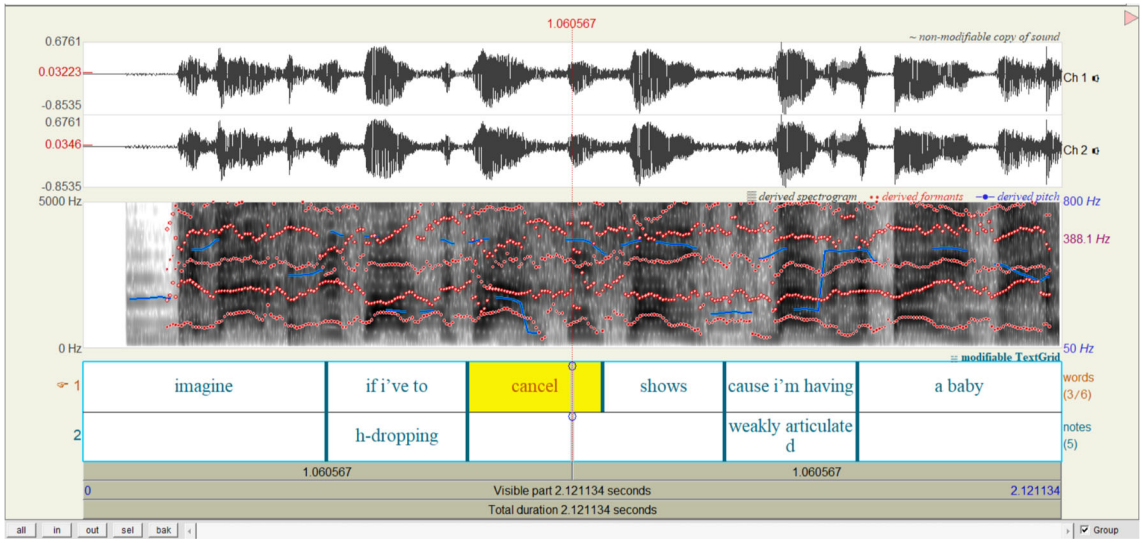


Figure 3. Spectrogram of Phrase 6: continuous formants across two extended stretches, "if I have to" and "cause I'm having", with no measurable boundaries within either stretch.

References

1. Abercrombie, D. Elements of General Phonetics. — Edinburgh : Edinburgh University Press, 1967. — 203 p.
2. Aiganatova, A. A. Challenges of teaching English pronunciation for Kazakh and Russian EFL students // Proceedings of the Student Scientific-Practical Conference. — Suleyman Demirel University, 2021. — P. 388–390. — URL: https://repository.sdu.edu.kz/server/api/core/bitstreams/defc70fc-08c6-4c11-8b0d9dd6f2d5d768/content
3. Benmezal, W., Bensemmane, F. Exploring students' difficulties with English connected speech and assimilation: A study at ENSB in Algiers // Educational and Didactic Research Review. — 2022. — Vol. 11, No. 3. — P. 443–462. — URL: https://www.asjp.cerist.dz/en/article/210531
4. Cauldwell, R. Greenhouse, Garden, Jungle [Electronic resource] // Speech in Action. — 2012. — URL: https://www.speechinaction.org/greenhouse-garden-jungle/ (accessed: 16.04.2026)
5. Crystal, D. English as a global language. — 2nd ed. — Cambridge : Cambridge University Press, 2003. — 212 p.
6. Dauer, R. M. Stress-timing and syllable-timing reanalyzed // Journal of Phonetics. — 1983. — Vol. 11, No. 1. — P. 51–62.
7. Field, J. Listening in the Language Classroom. — Cambridge : Cambridge University Press, 2008. — 366 p.
8. Ganong, W. F. Phonetic categorization in auditory word perception // Journal of Experimental Psychology: Human Perception and Performance. — 1980. — Vol. 6, No. 1. — P. 110–125.
9. Lindsey, G. English After RP: Standard British Pronunciation Today. — Palgrave Macmillan, 2019. — 153 p. — DOI: https://doi.org/10.1007/978-3-030-04357-5
10. Momen Omar, N., Pilus, Z. The effects of connected speech instruction on second or foreign language learners' perceptive skills and connected speech production: A systematic review of the literature (2000–2021) // International Journal of Learning, Teaching and Educational Research. — 2022. — Vol. 21, No. 9. — P. 389–414. — DOI: https://doi.org/10.26803/ijlter.21.9.22
11. Mulder, K., Brand, S., Boves, L., Ernestus, M. Processing reduced speech in the L1 and L2: A combined eye-tracking and ERP study // Language, Cognition and Neuroscience. — 2024. — Vol. 39, No. 5. — P. 527–551. — DOI: https://doi.org/10.1080/23273798.2024.2344162
12. Pike, K. L. The Intonation of American English. — Ann Arbor : University of Michigan Press, 1945. — 200 p.
13. Ramus, F., Nespor, M., Mehler, J. Correlates of linguistic rhythm in the speech signal // Cognition. — 1999. — Vol. 73, No. 3. — P. 265–292.
14. Teshaboyeva, N. Z., Khaydarova, M. The role of linking in English fluency development // Central Asian Journal of Multidisciplinary Research and Management Studies. — 2025. — Vol. 2, No. 12. — P. 177–179. — DOI: https://doi.org/10.5281/zenodo.17902031
15. van Heuven, W. J. B., Mandera, P., Keuleers, E., Brysbaert, M. SUBTLEX-UK: A new and improved word frequency database for British English // Quarterly Journal of Experimental Psychology. — 2014. — Vol. 67, No. 6. — P. 1176–1190. — DOI: https://doi.org/10.1080/17470218.2013.850521
16. Vandergrift, L., Tafaghodtari, M. H. Teaching L2 learners how to listen does make a difference: An empirical study // Language Learning. — 2010. — Vol. 60, No. 2. — P. 470–497.
17. Warren, R. M. Perceptual restoration of missing speech sounds // Science. — 1970. — Vol. 167, No. 3917. — P. 392–393.

18. Flege, J. E., Bohn, O.-S. The revised Speech Learning Model (SLM-r) // Second Language Speech Learning: Theoretical and Empirical Progress / ed. R. Wayland. — Cambridge : Cambridge University Press, 2021. — P. 3–83. — DOI: https://doi.org/10.1017/9781108886901.002

XARİCİ DİLİN ÖYRƏNİLMƏSİ VƏ TƏDRİSİNDƏ SÜNİ İNTELLEKT

SEVDA RAHIMOVA
Senior Teacher, Azerbaijan University of Foreign Languages

Açar sözlər: süni intellekt, xarici dil, xarici dil tədrisi, dil öyrənilməsi, generativ süni intellekt, ChatGPT, fərdiləşdirilmiş öyrənmə, dil bacarıqları, AI savadlılığı, rəqəmsal təhsil.

XXI əsrdə elm və texnologiyanın sürətli inkişafı insan fəaliyyətinin demək olar ki, bütün sahələrində ciddi dəyişikliklərə səbəb olmuşdur. Təhsil də bu transformasiya prosesindən kənarda qalmamışdır. İnformasiya-kommunikasiya texnologiyalarının tədris prosesinə daxil olması nəticəsində biliklərin əldə edilməsi, saxlanılması və ötürülməsi ilə yanaşı, öyrənmə və tədrisin təşkili formaları da dəyişmişdir. Bu dəyişikliklər xarici dil tədrisində daha aydın şəkildə müşahidə olunur. Xarici dilin öyrənilməsi artıq yalnız auditoriyada müəllimin izahını dinləmək və dərslikdə verilmiş tapşırıqları yerinə yetirməklə məhdudlaşmır. Elektron lüğətlər, onlayn təhsil platformaları, mobil tətbiqlər, videokonfrans vasitələri, avtomatik tərcümə sistemləri və müxtəlif rəqəmsal resurslar dil öyrənənlərə istənilən vaxt və məkanda dil üzərində işləmək imkanı verir.

Bu texnoloji inkişafın ən yeni mərhələsini süni intellekt təşkil edir. Süni intellekt uzun müddətdir müxtəlif sahələrdə istifadə edilsə də, generativ süni intellekt sistemlərinin geniş yayılması onun təhsil sahəsində tətbiqini tamamilə yeni səviyyəyə çıxarmışdır. Xüsusilə böyük dil modelləri əsasında fəaliyyət göstərən sistemlər insan dilini emal etmək, suallara cavab vermək, mətn yaratmaq, dialoq aparmaq, mətnləri təhlil etmək, tərcümə etmək və istifadəçinin tələbinə uyğun müxtəlif öyrənmə materialları hazırlamaq imkanına malikdir. Bu xüsusiyyətlər xarici dilin öyrənilməsi baxımından xüsusi əhəmiyyət daşıyır, çünki dil tədrisinin mərkəzində məhz dilin yaradılması, anlaşılması və kommunikativ məqsədlərlə istifadəsi dayanır.

Süni intellektin xarici dil tədrisinə daxil olması həm imkanlar, həm də problemlər baxımından yeni elmi müzakirələr doğurmuşdur. Əliyeva Vəfa “Süni intellekt və humanitar elmlər: müasir vəziyyət və perspektivlər” adlı məqaləsində qeyd edir: “Akademik ictimaiyyətdə rəqəmsal məzmunun dilə təsirinin dərəcəsi ilə bağlı çoxsaylı müzakirələr mövcuddur. J. Reed, V. Tsitskun və L. Baeva kimi bir çox alim rəqəmsallaşmanın dilin strukturuna mənfi təsir göstərdiyini və internet istifadəçilərinin ümumi savadlılığının azalmasına, onların idrak və ünsiyyət qabiliyyətlərinə təsir göstərdiyini və digər risklər daşıdığını qəbul edir. İnternetin meydana gəlməsi və inkişafı ilə ortaya çıxan yeni linqvistik ünsiyyətdə yeni bir nitq forması geniş yayılmış, redaktə edilməmiş və aşağı dərəcədə tənzimlənmişdir.

Bəzi filoloqlar orfoqrafiya və qrammatika standartlarının guya azalmasından şikayətlənsələr də, digərləri rəqəmsal ünsiyyətin linqvistik innovasiya və yaradıcılıq üçün yeni imkanlar açdığını iddia edirlər. E. Karavaeva və T. Kekeeva kimi bir qrup alim, veb resursların, əksinə, neologizmləri daxil etməklə, strukturunu yeniləməklə və böyük həcmdə məlumat ötürmək üçün daha rahat və qısa etməklə dili zənginləşdirdiyini iddia edirlər. Bu, dil öyrənmə və dil resursları ilə işləmək üçün innovativ metodların inkişafı üçün bir fürsət pəncərəsi açır.” [2] Bir tərəfdən, AI tələbəyə fərdi məşq imkanı verir, onun səhvlərini müəyyənləşdirir, müxtəlif səviyyələrə uyğun tapşırıqlar yaradır və dil öyrənməsini daha əlçatan edir. Digər tərəfdən isə AI-dən həddindən artıq istifadə tələbənin müstəqil düşünmə və yazma bacarıqlarına mənfi təsir göstərə, akademik dürüstlük məsələlərini kəskinləşdirə və yanlış məlumatların yayılmasına səbəb ola bilər.

Bu səbəbdən əsas məsələ süni intellektin xarici dil tədrisində istifadə edilib-edilməməsi deyil, onun hansı məqsədlərlə, hansı pedaqoji prinsiplərlə və hansı məhdudiyyətlər daxilində tətbiq

edilməsidir. Mövcud elmi araşdırmalar göstərir ki, AI-nin dil təhsilində rolu sürətlə genişlənməkdədir, lakin bu sahədə uzunmüddətli və empirik tədqiqatların sayı hələ məhduddur. Bu vəziyyət süni intellektin xarici dil öyrənilməsinə təsirinin daha sistemli şəkildə araşdırılmasını zəruri edir.

Xarici dil tədrisində texnologiyadan istifadə yeni hadisə deyil. Kompüter əsaslı dil öyrənmə proqramlarından başlayaraq mobil tətbiqlərə qədər müxtəlif texnologiyalar uzun müddətdir ki, dil öyrənənlərin fəaliyyətini dəstəkləyir. Lakin süni intellekti əvvəlki texnologiyalardan fərqləndirən əsas xüsusiyyət onun istifadəçi ilə qarşılıqlı əlaqədə olması və müəyyən mənada dinamik öyrənmə mühiti yaratmasıdır. Əvvəlcədən hazırlanmış elektron tapşırıq yalnız müəyyən cavabları qəbul edə bilər, generativ AI isə istifadəçinin cavabına uyğun yeni sual, izah, nümunə və tapşırıq yarada bilir.

Bu xüsusiyyət xarici dil tədrisinin fərdiləşdirilməsi baxımından böyük əhəmiyyət daşıyır. Ənənəvi sinifdə müəllim eyni dərsi müxtəlif səviyyəli tələbələrə təqdim etməli olur. Tələbələrin bir hissəsi mövzunu artıq bildiyi halda, digər hissəsi həmin mövzunu mənimsəməkdə çətinlik çəkə bilər. AI isə müəyyən dərəcədə bu fərqlərə uyğunlaşmağa imkan verir. Öyrənən öz səviyyəsini, məqsədini və çətinlik çəkdiyi sahəni müəyyənləşdirərək ona uyğun tapşırıq tələb edə bilər. Beləliklə, eyni texnologiyadan müxtəlif səviyyələrdə və müxtəlif məqsədlər üçün istifadə etmək mümkündür.

Süni intellektin digər mühüm xüsusiyyəti onun öyrənməni zaman və məkan baxımından daha çevik etməsidir. Tələbə yalnız dərs zamanı deyil, dərsdən sonra da dil praktikası apara bilər. Bu imkan xüsusilə xarici dil öyrənən şəxslər üçün əhəmiyyətlidir, çünki dil bacarıqlarının inkişafı müntəzəm təkrar və praktik fəaliyyət tələb edir. Sinif vaxtı məhdud olduğu halda, AI tələbəyə əlavə məşq üçün demək olar ki, fasiləsiz imkan yaradır.

Kohnke, Moorhouse və Zou (2023) generativ AI-nin dil tədrisində mətn yaratmaq, dil praktikası, ideyaların formalaşdırılması və müxtəlif tədris fəaliyyətlərinin dəstəklənməsi üçün istifadə edilə biləcəyini qeyd edirlər. Müəlliflərin fikrincə, AI-nin imkanlarından səmərəli istifadə üçün müəllimlər həm texnologiyanın imkanlarını, həm də onun məhdudiyyətlərini bilməlidirlər. Bu yanaşma göstərir ki, AI-nin təhsil prosesinə daxil olması müəllimin rolunu aradan qaldırmır, əksinə müəllimin texnologiya ilə işləmək bacarıqlarını daha da vacib edir.

Xarici dil öyrənilməsində əsas məqsədlərdən biri dörd fundamental dil bacarığının – dinləmə, danışma, oxu və yazının – inkişaf etdirilməsidir. Süni intellekt bu bacarıqların hər biri üçün müxtəlif imkanlar təqdim edir. Bununla belə, AI-nin bu sahələrdə effektivliyi onun necə tətbiq edilməsindən və digər pedaqoji fəaliyyətlərlə necə əlaqələndirilməsindən asılıdır.

Danışıq bacarığı xarici dil öyrənənlərin ən çox çətinlik çəkdiyi sahələrdən biridir. Tələbə qrammatika və söz ehtiyatı baxımından kifayət qədər biliyə malik olsa belə, real kommunikativ situasiyada danışmaqda çətinlik çəkə bilər. Bunun əsas səbəblərindən biri kifayət qədər danışıq praktikası imkanının olmamasıdır. Süni intellekt əsaslı chatbotlar və səsli sistemlər bu problemin müəyyən hissəsini həll edə bilər. Öyrənən AI ilə xarici dildə dialoq qura, müxtəlif situasiyaları canlandıra və öz ifadələrini təkrar-təkrar məşq edə bilər.

Daha mühüm məsələ isə tələbənin səhv etməkdən çəkinmədən danışa bilməsidir. İnsanlarla danışarkən səhv etməkdən utanmaq və ya sosial təzyiq hiss etmək mümkündür. AI ilə fərdi praktikada isə belə psixoloji maneələr nisbətən azalır.

Danışıq chatbotları ilə bağlı sistematik araşdırmalar göstərir ki, AI əsaslı danışıq vasitələri müəyyən hallarda öyrənənlərin danışıq bacarıqlarına, motivasiyasına və özünəinamına müsbət təsir göstərə bilər. Lakin bu nəticələr AI-nin real insan ünsiyyətini əvəz etdiyi anlamına gəlmir. Əksinə, AI ilə danışıq praktikası real kommunikasiya üçün hazırlıq mərhələsi kimi daha məqsədəuyğun hesab edilə bilər.

Yazı bacarığı sahəsində də süni intellekt mühüm imkanlar yaradır. Tələbə yazdığı mətni AI sisteminə təqdim etməklə qrammatik səhvləri, uyğun olmayan söz seçimini və bəzi üslub problemlərini müəyyənləşdirə bilər. Bundan əlavə, AI tələbəyə mətnin strukturunu qurmaq,

fikirləri sistemləşdirmək və alternativ ifadələr tapmaqda kömək edə bilər. Bu baxımdan AI formativ geribildirim vasitəsi kimi istifadə oluna bilər.

Lakin yazı bacarığında AI istifadəsi xüsusi diqqət tələb edir. Əgər tələbə bütün mətni AI-yə yazdırırsa, texnologiya onun yazı bacarığını inkişaf etdirmək əvəzinə bu bacarığın formalaşmasını əngəlləyə bilər. Xarici dil öyrənən tələbənin məqsədi mükəmməl mətn əldə etmək deyil, özü müstəqil şəkildə mətn yarada bilməkdir. Buna görə daha səmərəli yanaşma tələbənin ilkin mətni özü yazması, daha sonra AI-dən həmin mətni təhlil etməsini və səhvləri izah etməsini istəməsidir. Bu halda AI yazıçının əvəzi deyil, redaktor və öyrənmə köməkçisi funksiyasını yerinə yetirir.

Oxu bacarığının inkişafında da AI-nin imkanları genişdir. Süni intellekt müxtəlif mövzularda və müxtəlif çətinlik səviyyələrində mətnlər yarada bilər. Öyrənən mürəkkəb mətni sadələşdirilmiş formada tələb edə, bilmədiyi sözlərin izahını ala və mətn əsasında suallar hazırlada bilər. Bu imkanlar oxu fəaliyyətinin fərdiləşdirilməsinə şərait yaradır. Bununla belə, tələbənin yalnız AI tərəfindən yaradılan mətnlərlə məhdudlaşması düzgün olmazdı. Autentik məqalələr, kitablar, xəbər materialları və digər real dil nümunələri də tədris prosesində mühüm yer tutmalıdır.

Dinləmə bacarığı baxımından da texnologiyanın imkanları genişlənir. Süni intellekt əsaslı səs sistemləri müxtəlif mövzularda dialoqlar yaratmağa və nitqin müxtəlif sürətlərdə təqdim edilməsinə imkan verir. Tələbə əvvəlcə daha yavaş nitqlə, sonra isə təbii danışıq sürəti ilə işləyə bilər. Bu, xüsusilə xarici dilin danışıldığı mühitdə yaşamayan tələbələr üçün faydalı ola bilər.

Qrammatika da xarici dil tədrisinin mühüm komponentlərindən biridir. Ənənəvi dil tədrisində qrammatik qaydalar çox vaxt müəllim tərəfindən izah edilir və tələbələr müxtəlif tapşırıqlar vasitəsilə həmin qaydaları möhkəmləndirirlər. Süni intellekt bu prosesi daha interaktiv edə bilər. Tələbə konkret cümlənin niyə səhv olduğunu soruşa və qaydanın müxtəlif kontekstlərdə necə işlədiyini öyrənə bilər.

Bu yanaşmanın üstünlüyü ondadır ki, tələbə yalnız “düzgün” və ya “səhv” cavabı deyil, səhvin səbəbini də öyrənmək imkanı əldə edir. Bununla belə, AI-nin qrammatik izahları avtomatik olaraq doğru qəbul edilməməlidir. Dilin bəzi mürəkkəb və kontekstdən asılı xüsusiyyətlərini süni intellekt yanlış izah edə bilər. Bu səbəbdən müəllim nəzarəti və etibarlı dil mənbələri mühüm olaraq qalır.

Lüğət öyrənilməsi də AI-nin faydalı ola biləcəyi sahələrdən biridir. Xarici dil öyrənən şəxs yeni sözün mənasını öyrənməklə kifayətlənməməli, həmin sözün hansı kontekstdə işlədiyini də bilməlidir. AI sözün sinonimlərini, antonimlərini, kollokasiyalarını və nümunə cümlələrini təqdim edə bilər. Hətta tələbənin artıq öyrəndiyi sözlər əsasında yeni məşqlər yaratmaq mümkündür.

Beləliklə, AI lüğət öyrənməsini mexaniki əzbərləmə prosesindən daha interaktiv fəaliyyətə çevirə bilər. Tələbə yeni sözləri müxtəlif situasiyalarda istifadə etməklə onların aktiv lüğətə keçməsinə nail ola bilər.

Fərdiləşdirilmiş öyrənmə müasir təhsilin digər əsas istiqamətlərindən biridir. Xarici dil tədrisində fərdiləşdirmənin əhəmiyyəti xüsusilə böyükdür, çünki öyrənənlərin dil səviyyələri, öyrənmə sürətləri, motivasiyaları və ehtiyacları bir-birindən əhəmiyyətli dərəcədə fərqlənir.

Süni intellekt müəyyən dərəcədə bu fərqləri nəzərə almağa imkan verir. Tələbə öz dil səviyyəsinə uyğun mətn, tapşırıq və izah tələb edə bilər. Məsələn, eyni mövzu A2 səviyyəsində sadə sözlərlə, B1 səviyyəsində orta mürəkkəblikdə və C1 səviyyəsində akademik üslubda təqdim edilə bilər.

Bu xüsusiyyət öyrənənin öz sürətində inkişaf etməsinə imkan yaradır. Tələbə başa düşmədiyi mövzunu dəfələrlə izah etdirə və əlavə nümunələr tələb edə bilər. Beləliklə, fərdi öyrənmə imkanları genişlənir.

Lakin burada mühüm bir məqam nəzərə alınmalıdır. Fərdiləşdirmə yalnız texnologiyanın təqdim etdiyi məlumatlara əsaslanmamalıdır. Müəllim tələbənin motivasiyasını, emosional vəziyyətini, öyrənmə vərdişlərini və sosial xüsusiyyətlərini də nəzərə almalıdır. Buna görə AI əsaslı fərdiləşdirmə müəllim tərəfindən həyata keçirilən pedaqoji fərdiləşdirməni tamamlamalıdır.

Süni intellektin təhsildə geniş yayılması müəllimin gələcək rolu ilə bağlı mühüm suallar yaratmışdır. Texnologiyanın böyük həcmdə informasiya təqdim edə bilməsi bəzən müəllimin əvvəlki funksiyalarının bir hissəsinin avtomatlaşdırılacağı fikrini doğurur. Lakin müəllimin fəaliyyətini yalnız informasiya ötürülməsi ilə məhdudlaşdırmaq düzgün deyil.

Müəllim təlim məqsədlərini müəyyənləşdirir, tələbələrin ehtiyaclarını qiymətləndirir, onların motivasiyasını formalaşdırır, sinifdə qarşılıqlı əlaqəni təşkil edir və öyrənmə prosesinə emosional və sosial dəstək verir. Bu funksiyalar sadəcə informasiya emalı ilə əvəz edilə bilməz.

Bu baxımdan süni intellekt müəllimin rolunu aradan qaldırmaqdan çox, onun xarakterini dəyişir. Müəllim daha çox mentor, fasilitator, təlim dizayneri və texnologiyanın pedaqoji istifadəsini idarə edən mütəxəssis rolunda çıxış edə bilər. Müəllim AI-dən alınan materialları seçməli, onların düzgünlüyünü yoxlamalı və tələbələrin səviyyəsinə uyğunlaşdırmalıdır.

Zawacki-Richter və həmkarlarının (2019) ali təhsildə süni intellektlə bağlı apardıqları sistematik araşdırma göstərmişdir ki, AI üzrə tədqiqatlarda pedaqoqların perspektivi kifayət qədər diqqət mərkəzində olmayıb. Bu nəticə xarici dil tədrisi üçün də mühüm əhəmiyyət daşıyır. AI-nin təhsildə effektiv tətbiqi müəllimin prosesdən kənarlaşdırılması ilə deyil, müəllimin texnologiyanın tətbiqinə fəal şəkildə cəlb edilməsi ilə mümkündür.

Xarici dil öyrənməsində tələbənin müstəqilliyi xüsusi əhəmiyyət daşıyır. Dil yalnız dərs zamanı öyrənilən bilik deyil; onun inkişafı davamlı məşq tələb edir. AI bu baxımdan tələbənin müstəqil öyrənmə imkanlarını genişləndirə bilər.

Tələbə müəllimdən asılı olmadan müəyyən mövzunu təkrar edə, suallar verə və əlavə məşqlər əldə edə bilər. Bu, öyrənənin öz təlim prosesinə nəzarətini artırır. Bununla yanaşı, müstəqillik AI tərəfindən təqdim edilən bütün məlumatların qəbul edilməsi demək deyil.

Əksinə, AI dövründə tənqidi düşünmə daha böyük əhəmiyyət qazanır. Tələbə AI-nin cavabını qiymətləndirməli, onun düzgünlüyünü digər mənbələrlə müqayisə etməli və lazım gəldikdə müəllimə müraciət etməlidir. Bu baxımdan müasir dil öyrənən şəxsin yalnız dil kompetensiyasına deyil, həm də rəqəmsal və AI savadlılığına ehtiyacı vardır.

Süni intellektin xarici dil tədrisində yaratdığı imkanlarla yanaşı, ciddi problemləri də mövcuddur. Əsas problemlərdən biri AI tərəfindən yaradılan məlumatların etibarlılığıdır. Generativ süni intellekt sistemləri bəzən yanlış məlumatı çox inandırıcı formada təqdim edə bilər. Dil tədrisində bu, xüsusilə təhlükəlidir, çünki tələbə yanlış qrammatik qaydanı və ya qeyri-təbii ifadəni düzgün hesab edə bilər.

Digər mühüm problem akademik dürüstlüklə bağlıdır. Tələbə AI-dən istifadə etməklə esse, təqdimat və digər yazılı tapşırıqları özü yazmadan hazırlaya bilər. Bu vəziyyət ənənəvi qiymətləndirmə sistemlərinin effektivliyini azaldır. Lakin AI-ni tamamilə qadağan etmək real və uzunmüddətli həll yolu deyil. Bunun əvəzinə tələbələrə AI-dən akademik və etik şəkildə istifadə qaydaları öyrədilməlidir.

AI-dən həddindən artıq istifadə də problem yarada bilər. Əgər tələbə hər bir cümləni, hər bir fikri və hər bir yazını AI-yə həvalə edirsə, onun müstəqil dil yaratmaq bacarığı zəifləyə bilər. Xarici dil öyrənməsində müəyyən çətinliklərin yaşanması təbii və faydalı prosesdir. Tələbə özü düşünməli, söz axtarmalı, səhv etməli və həmin səhvlərdən nəticə çıxarmalıdır.

Məxfilik məsələsi də xüsusi diqqət tələb edir. Tələbələr AI sistemlərinə şəxsi məlumatlar və akademik işlər daxil edə bilərlər. Bu məlumatların necə saxlanıldığı və emal edildiyi ilə bağlı şüurlu davranış tələb olunur. UNESCO generativ AI-nin təhsildə istifadəsinə dair tövsiyələrində insan mərkəzli yanaşmanı, məlumatların qorunmasını, bərabərliyi və etik prinsipləri xüsusi olaraq vurğulayır.

Bundan əlavə, rəqəmsal bərabərsizlik problemi mövcuddur. Bütün tələbələrin müasir cihazlara, yüksək sürətli internetə və eyni AI alətlərinə çıxışı olmaya bilər. Bu vəziyyət texnologiyanın təhsildə tətbiqi zamanı yeni bərabərsizlik formalarının yaranmasına səbəb ola bilər.

Süni intellektin xarici dil tədrisində tətbiqi Azərbaycan təhsil sistemi üçün də aktual məsələlərdən biridir. Xarici dillərin, xüsusilə ingilis dilinin öyrənilməsinə olan tələbatın artması rəqəmsal texnologiyalardan istifadənin əhəmiyyətini artırır. AI əsaslı vasitələr Azərbaycanlı tələbələrə dərsdənkənar danışıq praktikası aparmaq, qrammatik və leksik problemləri araşdırmaq, yazılarını təkmilləşdirmək və müxtəlif mövzularda xarici dildə ünsiyyət qurmaq imkanları verə bilər.

Azərbaycan dili ilə xarici dillər arasında struktur və qrammatik fərqlər olduğuna görə AI-dən istifadə zamanı müqayisəli izahların keyfiyyətinə xüsusi diqqət yetirilməlidir. Məsələn, ingilis dilində zaman sistemi, artikllar, söz sırası və bəzi sintaktik konstruksiyalar Azərbaycan dili ilə müqayisədə fərqli xüsusiyyətlərə malikdir. AI bu fərqləri izah etməkdə faydalı ola bilər, lakin onun təqdim etdiyi izahlar müəllim və etibarlı dil mənbələri tərəfindən yoxlanmalıdır.

Azərbaycan universitetlərində xarici dil müəllimlərinin AI savadlılığının artırılması mühüm istiqamətlərdən biri ola bilər. Müəllim yalnız AI alətlərindən istifadə etməyi deyil, həm də tələbələrə onların düzgün və etik istifadəsini öyrətməlidir. Eyni zamanda, qiymətləndirmə sistemlərinin AI dövrünün reallıqlarına uyğunlaşdırılması vacibdir.

Süni intellekt texnologiyalarının inkişafı davam etdikcə onların xarici dil tədrisində tətbiq sahələrinin də genişlənəcəyi gözlənilir. Gələcəkdə daha təkmil səsli sistemlər, real vaxt rejimində danışıq analizi, fərdi öyrənmə platformaları və daha dəqiq dil geribildirimi xarici dil tədrisinin ayrılmaz hissəsinə çevrilə bilər.

Lakin texnologiyanın inkişafı ilə yanaşı pedaqoji yanaşmalar da dəyişməlidir. Əsas diqqət “AI nə edə bilər?” sualından “AI öyrənmə prosesinə hansı pedaqoji dəyəri əlavə edir?” sualına yönəlməlidir.

Bu baxımdan gələcəyin xarici dil tədrisini üç əsas komponentin əməkdaşlığı kimi təsəvvür etmək mümkündür: müəllim, tələbə və süni intellekt. Müəllim pedaqoji məqsədləri müəyyən edir və prosesi istiqamətləndirir; AI fərdi məşq, geribildirim və əlavə resurslar təqdim edir; tələbə isə öyrənmə prosesinin aktiv iştirakçısı kimi çıxış edir.

Belə modeldə texnologiya insan münasibətlərini əvəz etmir. Əksinə, müəllimin vaxtını təkrarlanan fəaliyyətlərdən azad etməklə onun tələbələrlə daha çox fərdi və yaradıcı iş aparmasına şərait yarada bilər. Eyni zamanda tələbə daha çox müstəqillik əldə edir, lakin bu müstəqillik tənqidi düşünmə və məsuliyyətlə müşayiət olunur.

Nəticə

Süni intellekt müasir xarici dil tədrisinin inkişafında mühüm rol oynamağa başlayan texnologiyalardan biridir. Onun imkanları xarici dil öyrənməsinin müxtəlif mərhələlərini əhatə edir və xüsusilə fərdiləşdirilmiş öyrənmə, danışıq praktikası, yazı bacarıqları, qrammatika, lüğət, oxu və dinləmə fəaliyyətləri baxımından yeni imkanlar yaradır. Generativ AI-nin əsas üstünlüyü onun öyrənənlə qarşılıqlı əlaqəyə girə bilməsi və konkret ehtiyaca uyğun məzmun yarada bilməsidir.

Bununla yanaşı, süni intellektin təhsilə daxil olması avtomatik olaraq daha keyfiyyətli öyrənmə demək deyil. Texnologiyanın pedaqoji məqsədsiz və nəzarətsiz istifadəsi yanlış məlumatların mənimsənilməsinə, akademik dürüstlüyün pozulmasına, texnologiyadan asılılığın artmasına və real insan ünsiyyətinin azalmasına səbəb ola bilər. Bu səbəbdən AI-nin xarici dil tədrisində tətbiqi həm texnoloji, həm də pedaqoji və etik baxımdan düşünülmüş şəkildə həyata keçirilməlidir.

Məqalənin təhlili göstərir ki, süni intellektin ən səmərəli istifadəsi onun müəllimi əvəz etməsi deyil, müəllimin fəaliyyətini tamamlaması və tələbənin müstəqil öyrənmə imkanlarını genişləndirməsidir. Müəllim pedaqoji istiqamət və nəzarəti təmin etməli, tələbə aktiv öyrənən olaraq qalmalı, süni intellekt isə fərdi məşq və geribildirim vasitəsi kimi istifadə edilməlidir.

Bu baxımdan gələcəyin xarici dil tədrisi insan və texnologiya arasında rəqabət üzərində deyil, əməkdaşlıq üzərində qurulmalıdır. Süni intellektin əsas dəyəri insan müəllimini əvəz etməkdə deyil, onun imkanlarını genişləndirməkdə və öyrənənə daha çevik, fərdi və interaktiv təlim mühiti yaratmaqda ifadə olunur. Xarici dil tədrisinin gələcək inkişafı üçün əsas istiqamət texnoloji

yenilikləri pedaqoji prinsiplərlə əlaqələndirmək, müəllimlərin AI savadlılığını artırmaq, tələbələrdə tənqidi düşünmə və rəqəmsal kompetensiya formalaşdırmaq və real insan ünsiyyətini təlim prosesinin mərkəzində saxlamaqdır.

Ədəbiyyat

1. Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868
2. Алиева В. Лингвистические трансформации английского языка в условиях цифровизации. "Süni intellekt və humanitar elmlər: müasir vəziyyət və perspektivlər" Beynəlxalq elmi konfransın materialları. – Bakı: ADU, 2025 – 464 s.
3. Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
4. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://doi.org/10.1186/s41239-019-0171-0
5. Godwin-Jones, R. (2023). Emerging technologies: Generative AI and language learning. *Language Learning & Technology*.
6. Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.
7. Chapelle, C. A. (2001). *Computer Applications in Second Language Acquisition*. Cambridge University Press.
8. Levy, M., & Stockwell, G. (2006). *CALL Dimensions: Options and Issues in Computer-Assisted Language Learning*. Lawrence Erlbaum Associates.
9. Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31*(2), 57–71.

Psychological Sciences

Exploring the interplay between intrinsic and extrinsic motivation in student learning

Ramazanova K.
Sheki branch of Azerbaijan State Pedagogical University, Psychology teacher, Azerbaijan, Sheki, https://orcid.org/0009-0004-5225-3779

ABSTRACT

This article analyzes the psychological and pedagogical essence of motivation in the learning process of students, the impact of types of motivation (intrinsic and extrinsic motivation) on student activity, and the strategies for forming motivation in a modern educational environment on a scientific basis. The theoretical framework of the study is based on the Self-Determination Theory (SDT) developed by Edward Deci and Richard Ryan. This theory emphasizes that the quality of motivation depends on the satisfaction of the student's three fundamental psychological needs - autonomy, competence, and relatedness. The article shows that intrinsic motivation plays a significant role in the development of processes such as deep learning, creativity, and sustained effort by the student, while extrinsic motivation performs the function of initiating, regulating behavior, and engaging in learning at the initial stages.

The impact of motivation on the quality of learning has been widely studied in Azerbaijani pedagogical literature (Taghiyeva, 2023; Haziyeva, 2024; Gahramanli, 2025, etc.) and it has been confirmed that student success directly depends on the level of motivation. Local studies emphasize that motivation is formed not only by psychological, but also by social, pedagogical and organizational factors. A comparison of international approaches with local experience shows that an effective motivation model requires the creation of a learning environment that takes into account the individual potential, interests, learning style and psychological needs of the student. The article ultimately proves that motivation is not only a component of learning, but also the main psychological mechanism that determines all stages of learning. Intrinsic motivation is considered the most powerful factor leading to personality development, lifelong learning and sustainable success in learning, while extrinsic motivation, when properly applied, plays the role of a bridge for the student's motivation to transition to an intrinsic form.

KEYWORDS: intrinsic motivation, extrinsic motivation, learning, training, achievement

Modern education systems demand learning strategies that take into account a student's personality, individual characteristics, interests, and inclinations. Today, the primary objective of the instructional process is not merely imparting knowledge, but also fostering student engagement, uncovering their potential, and building sustainable learning motivation. Within the learning process, motivation acts as the central psychological factor directing student activity and shaping their attitude toward learning; it dictates their interest in lessons, resilience against challenges, learning speed, and overall achievement levels.

Motivation is broadly divided into two main categories: intrinsic motivation, stemming from the inherent interest and satisfaction derived from the learning process itself, and extrinsic motivation, formed through external factors such as grades, rewards, or social approval. Studying these two types of motivation is crucial for understanding the essence of the educational process and enhancing the effectiveness of pedagogical methods applied by teachers.

In recent years, Ryan and Deci's Self-Determination Theory has played a particularly significant role in explaining the psychological foundations of motivation. According to this theory, sustained and high-quality human behavior relies on satisfying three fundamental psychological needs: autonomy, competence, and relatedness. Meeting these needs within an educational environment creates conditions for students to become more active, engaged, and persistent in their learning process.

Azerbaijani pedagogical research also emphasizes the role of motivation in learning, with local authors demonstrating that motivation impacts both the quality of instruction and the student's self-confidence and active participation. From this perspective, examining the essence of motivation on a scientific-theoretical basis and determining its implementation mechanisms in learning is one of the crucial imperatives of the modern education system.

The objective of this article is to comparatively analyze students' intrinsic and extrinsic types of motivation, investigate their roles in the instructional process, and draw upon existing theories to demonstrate the effective application of motivation in education.

The success of the educational process relies on new pedagogical thinking. One of the foundational pillars of this new pedagogical thinking is creating suitable conditions and motivation for the student to want to study well. A student's desire to excel is conditioned by various factors within a psychological context, yet its core remains connected to the sphere of motivation—specifically learning motivation. The lifeblood of this domain develops, first and foremost, through learning interests. A student must want to learn and be interested in the lesson. School experience reveals many capable students who fail to achieve good results simply due to a lack of motivation.

The modern educational approach considers it essential to account for a student's personal characteristics, abilities, range of interests, and individual learning style. The primary goal in the learning process is to develop these traits, uncover the student's potential, and enrich it further. In such a model, a student's lagging behind is not framed as a personal fault; instead, the underlying cause of the difficulty is investigated, and appropriate support is provided without undermining the student's personality. Along with this, the student's interest in learning is heightened, and conditions are created for them to demonstrate their abilities and find their path in areas of interest. As a result, the student does not grow tired or bored during class; rather, they experience the learning process as one of the most rewarding and motivating aspects of their life.

The nature of learning motivation and personality traits are essentially key indicators of the quality of education. For successful learning, the motivation factor is no less important than intellect. Recognizing the high significance of motives for achievement led to the substantiation of the principle of motivation in the instructional process. In various categorizations of motivation, psychologists express positions regarding essential issues, specifically the relationship between the energy and content aspects of motivation (Taghiyeva, S. 2023).

Learning motivation is carried out by the individual as a conscious activity. These influences are tied to the characteristics of a person's life relationships, desires, and interests, which reflect the nature of their motivation and learning motives. The term "motivation" was first used as a technical term by A. Schopenhauer in his essay "On the Fourfold Root of the Principle of Sufficient Reason," and was subsequently applied to explain the causes of human and animal behavior.

According to L. I. Bozhovich, motives can be objects of the external world, mental representations, ideas, feelings, and emotions. At the same time, motive and motivation manifest as a complex mechanism through which an individual coordinates external and internal factors of behavior, determining the emergence, direction, and implementation paths of corresponding forms of activity (Mansimova, K. G. 2023).

According to E. G. Izotova, motivation is the aggregate of motives. Motivation is one of the fundamental problems in both domestic and foreign psychology. Its importance for the

development of modern psychology is tied to the analysis of the sources of human activity and the driving forces behind an individual's actions and behavior (Izotova, E. G. 2014).

Educational activity carries different meanings for different students. Therefore, identifying a student's learning motivation and the personal meaning of education in every concrete situation plays a decisive role in determining pedagogical and psychological intervention measures for school children. During the learning process, students are influenced not by a single motive, but by a system of motives. Practice shows that all motivational relations giving an active and directed character to learning activity enter this structure and determine its content-semantic characteristics. This system of motives constitutes learning motivation, which is characterized by both stability and dynamism.

Generally, there are three types of activity sources: internal, external, and personal. The internal sources of learning motivation include cognitive and social needs (the desire for socially approved actions and achievements). The external sources of learning motivation are determined by the student's living conditions, which encompass demands, expectations, and opportunities. Demands relate to the need to adhere to social norms of behavior, communication, and activity. Expectations characterize society's attitude toward education as a norm of behavior accepted by the individual, enabling them to overcome challenges associated with carrying out educational activities (Hasanov, I. 2023).

Motivation is the heart of the learning process. A student attending class, completing assignments, and striving to gain knowledge is the result of a specific force existing either within or outside of them. Motivation in education is divided into two primary categories: Intrinsic Motivation and Extrinsic Motivation. The impact of these two forces on a student's quality of learning, academic achievement, and even future life success is immense.

Intrinsic motivation occurs when a student engages in an activity for its own sake, independent of any external reward or punishment. In this case, the learning process itself acts as a reward. The student learns the lesson, topic, or task because it is interesting, deriving pleasure directly from the process. Intrinsic motivation stems from the student's need to enhance their knowledge and skills, overcome challenges, and achieve self-development; they simply want to "become better" (Gahramanli, N. J. 2025).

Self-Determination Theory (SDT) is a theory of motivation formulated in the late 1970s and early 1980s by Edward Deci and Richard Ryan. Self-Determination Theory is a humanistic-cognitive approach that explains the source, persistence, and quality of human behavior. The central idea of the theory is that "human nature is inherently active, oriented toward self-development, and directed at realizing its potential in interaction with the social environment." However, for this to occur, the environment must satisfy specific psychological needs. The theory explains the quality of both intrinsic and extrinsic motivation and measures the degree to which behavior is self-endorsed.

According to Self-Determination Theory, people are innately driven by three basic psychological needs: competence, relatedness, and autonomy. Autonomy is the need for a person to regulate their own behavior based on choice, control, and volition, expressing internal agency through questions like "What do I want?" and "How do I want to do it?". Competence is the need for a person to feel effective, capable, and skilled; people exert effort to experience a sense of accomplishment. Relatedness is the need for social connectedness, acceptance, care, and belonging; when individuals feel valued by others, their motivation grows stronger (Ryan, R. M., & Deci, E. L. 2017). Intrinsic motivation specifically feeds on the student's desire to make choices and feel capable.

Research conducted by Deci and Ryan in the 1970s revealed an interesting phenomenon known as the "Overjustification Effect." This occurs when an external reward (such as money or grades) is introduced for an activity originally driven by intrinsic motivation, resulting in a decrease in the

student's intrinsic interest in that activity. For instance, if a student who genuinely loves drawing begins receiving money for every drawing they create, after a while, they will draw solely to earn money rather than for pure enjoyment. If the financial reward is discontinued, the student's interest in drawing will be lower than it was initially. This illustrates how an external reward can "overshadow" intrinsic interest.

Intrinsically motivated students exert greater effort, do not shy away from challenging tasks, and retain what they learn for a long duration. Instead of simply memorizing information, they strive to truly comprehend it. Individuals who perform an activity for pure enjoyment take more risks and seek non-standard solutions, which significantly enhances their creativity. Students who learn through their own choices feel less pressure, possess higher self-esteem, and are happier at school.

Extrinsic motivation refers to a student performing an activity to achieve an external outcome (such as rewards, grades, praise) or to avoid punishment. According to Self-Determination Theory (SDT), extrinsic motivation is not a single uniform category; rather, it exists across four distinct levels that reflect how closely the "regulation" aligns with an internal state:

* External Regulation: The student learns solely to receive a reward or avoid punishment.
* Introjected Regulation: The student acts under internal pressures such as "guilt" or "pride." Although the drive for the action remains external, the feeling is internal. The student learns to avoid losing the teacher's approval or to avoid feeling bad about themselves.
* Identified Regulation: The student internalizes the value of the activity rather than the activity itself. They realize that this activity helps them achieve personal goals. For example, a student studies biology because they want to become a doctor in the future, even if the subject itself is not particularly interesting to them.
* Integrated Regulation: The activity completely aligns with the student's personal values and identity. This is the form of extrinsic motivation closest to intrinsic motivation (Ryan & Deci, 2017).

Extrinsic motivation is crucial when a student needs an initial push to begin an activity or when they must navigate topics that lack inherent interest. External rewards can quickly initiate student action and prove effective for completing monotonous or uninteresting tasks. Furthermore, they help spark initial engagement when introducing a new topic (Jabrayilova, R. N., 2024).

However, the drawbacks of extrinsic motivation are evident: as soon as the student completes the task or the reward stops, the motivation vanishes. The primary objective becomes acquiring the reward rather than genuine learning. Over-reliance on reward and punishment systems can reduce students' sense of control over their actions. Ultimately, students tend to choose only easy tasks to minimize the risk of failure, as their overarching goal is simply to achieve a high result or reward.

In an ideal learning process, the main goal of teachers and parents is to guide the student gradually toward intrinsic motivation by leveraging extrinsic motivation—specifically "identified" and "integrated" regulations. This process is known as the internalization of motivation. In this way, even if a student does not find a lesson inherently interesting, they study because they comprehend that it is a necessary instrument to achieve their personal goals. This elevates the meaning of learning and ensures long-term retention.

Conclusion

The conducted analyses demonstrate that motivation is the primary psychological element driving both the initiation and sustainability of the learning process. Students' levels of intrinsic and extrinsic motivation directly impact their learning behavior, cognitive engagement, problem-solving skills, and academic achievements. Intrinsic motivation gives meaning to learning activities, enhances interest and cognitive pleasure, and transforms learning into a process of personal growth. This form of motivation empowers students to overcome challenges while fostering key 21st-century skills such as creativity, initiative, and critical thinking.

At the same time, extrinsic motivation plays a significant role in the early stages of learning and when mastering topics that students find uninteresting. However, research (Deci & Ryan, 2000;

Mansimova, 2023; Suleymanli, 2024) indicates that extrinsic motivation does not sustain long-term learning on its own; if relied upon exclusively, it can reduce student self-confidence and cause the learning process to become mechanical. Therefore, extrinsic motivation should serve as a functional tool designed to support the transition toward intrinsic motivation.

The analysis of Azerbaijani pedagogical literature shows that modern pedagogy prioritizes building a supportive, relational learning environment focused on the student's personality and social-emotional needs. Local studies emphasize that key factors in boosting motivation include:

* Positive teacher-student relationships
* Connecting lesson topics to real-world applications
* Cultivating independent learning skills
* Implementing differentiated instruction

Ultimately, the most effective motivational strategy is creating an educational environment that satisfies students' basic psychological needs, provides freedom of choice, instills a sense of achievement, and fosters social acceptance. When intrinsic and extrinsic motivation are balanced effectively, students engage in the learning process not merely as passive participants, but as active authors of their own personal development—resulting in both higher academic success and greater social-emotional well-being.

REFERENCES

1. Tağıyeva, S. (2023). Təlim-tərbiyə prosesində motivasiyanın yüksəldilməsi yolları. Məktəbəqədər və İbtidai Təhsil Jurnalı. https://ppe-journal.edu.az/frontend/uploads/files/Sevda%20Taghiyeva.pdf
2. Həziyeva, M. Z. (2024). Ali təhsil müəssisələrində tələbələrin təlimə motivasiyası. AR Təhsil İnstitutu Elmi Jurnal. https://journals.arti.edu.az/uploads/f_20241121081530789276137.pdf
3. Cəbrayılova, R. N. (2024). Ali məktəblərdə tələbələrin təlim fəaliyyətinin motivasiyasının müasir aspektləri. AR Təhsil İnstitutu Elmi Jurnal. https://journals.arti.edu.az/uploads/f_20240801080846987008885.pdf
4. Qəhrəmanlı, N. C. (2025). Təhsildə nailiyyət motivasiyasının əsas istiqamətləri. AR Təhsil İnstitutu Elmi Jurnal. https://journals.arti.edu.az/uploads/f_20250109071401106972751O.pdf
5. Mənsimova, K. Q. (2023). Nailiyyətin motivasiyası, özünütənzimləmə və təhsil müvəffəqiyyəti. Azərbaycan Elmi-Metodik Mərkəzi. https://aem.az/nalyyetn-motvasyas-ozunutenzmleme-ve-tehsl-muveffeqyyet
6. Həsənov, İ. (2022). Motivasiya: fəal dərsin mərhələləri kimi. Təhsilin Ali İdarəetmə Məktəbi. https://taim.edu.az/motivasiya
7. Ryan, R. M., & Deci, E. L. (2017). Self-Determination Theory: Basic psychological needs in motivation, development, and wellness. https://doi.org/10.1521/978.14625/28806 https://www.guilford.com/excerpts/ryan.pdf?t=1
8. Изотова, Е. Г. Мотивация учебной деятельности студентов вузов: особен- ности формирования и развития [Текст] / Е.Г. Изотова //Реализация стандартов второго поколения в школе: проблемы и перспективы: сборник научных статей третьей всероссийской интернет-конференции (октябрь – декабрь 2014 г.) / под науч. ред. Е. В. Карповой. – Ярославль: Изд-во ЯГПУ, 2014. – 190 с. – С. 54–59.

FATHER'S INFLUENCE ON THE FORMATION OF A CHILD'S GENDER IDENTITY

Aliyeva Khanym Ibrahim
Baku State University, Department of Psychology, PhD Student, Baku, Azerbaijan

Abstract: This research examines the figure of the father as a key factor in the development of a child's gender identity and sexual attitudes. The father contributes to the formation of a sense of security and self-confidence. He exerts not only behavioral influence but also a deep emotional impact on the development of the child's sexuality. The work emphasizes the importance of the father's role in the harmonious development of the child.
A father's absence or emotional coldness can hinder the development of gender role behavior and negatively impact a child's self-esteem. Conversely, hypercriticism or aggressiveness on the part of the father can evoke feelings of fear and emotional detachment, influencing the nature of future interpersonal relationships. Conversely, in families where the father demonstrates respect, support, and emotional engagement, children are more likely to develop a stable gender identity and a positive attitude toward their gender role. A father's active involvement in upbringing promotes a more harmonious development of gender roles and reduces the likelihood of stereotypical distortions.
Keywords: gender identity, the role of the father, the Jungian approach

ВЛИЯНИЕ ОТЦА НА ФОРМИРОВАНИЕ ГЕНДЕРНОЙ ИДЕНТИЧНОСТИ РЕБЁНКА

Алиева Х.И.
Бакинский Государственный Университет,
Кафедра Психологии, Докторант
Баку, Азербайджан

Аннотация: В данном исследовании рассматривается роль отца как ключевого фактора в формировании гендерной идентичности и сексуальных установок ребенка. Отец способствует формированию чувства безопасности и уверенности в себе. Он оказывает не только поведенческое, но и глубокое эмоциональное влияние на развитие сексуальности ребенка. Работа подчеркивает важность роли отца в гармоничном развитии ребенка.
Отсутствие отца или его эмоциональная холодность могут препятствовать формированию гендерного поведения и негативно влиять на самооценку ребенка. И наоборот, чрезмерная критика или агрессивность со стороны отца могут вызывать чувство страха и эмоциональную отстраненность, влияя на характер будущих межличностных отношений. Напротив, в семьях, где отец проявляет уважение, поддержку и эмоциональную вовлеченность, дети с большей вероятностью формируют стабильную гендерную идентичность и позитивное отношение к своей гендерной роли. Активное участие отца в воспитании способствует более гармоничному формированию гендерных ролей и снижает вероятность стереотипных искажений.
Ключевые слова: гендерная идентичность, роль отца, юнгианский подход

Введение: Роль отца является важным и многогранным фактором, оказывающим влияние на различные аспекты развития ребёнка. Согласно работам Майкла Лэмба, отцы вносят уникальный вклад в эмоциональное развитие детей, прежде всего посредством активных и стимулирующих форм взаимодействия, таких как игра, способствующих развитию

навыков эмоциональной регуляции. Кроме того, устойчивое участие отца в жизни ребёнка связано с более высоким уровнем социальной компетентности, что способствует успешному установлению и поддержанию межличностных отношений.

Одной из ключевых функций отца является выполнение роли модели гендерного поведения. В процессе наблюдения и взаимодействия дети усваивают нормы поведения, социальные роли и элементы личностной идентичности. Наряду с этим вовлечённость отца оказывает положительное влияние на когнитивное развитие ребёнка, включая формирование языковых навыков и способности к решению задач, особенно в случаях его активного участия в повседневной жизни семьи [1].

В сборнике Handbook of Fathers and Child Development подчёркивается, что участие отца следует рассматривать в более широком контексте межличностных и культурных отношений, где гендерные роли формируются под влиянием семейной динамики, триадных взаимодействий и социокультурных норм. Влияние отца проявляется не только непосредственно через взаимодействие с ребёнком, но и косвенно – через формирование семейной системы, а также поддержание или трансформацию существующих гендерных ожиданий.

Авторы также отмечают, что разнообразие отцовских ролей в различных культурах свидетельствует о том, что, несмотря на различия в формах проявления гендерного моделирования, отец остаётся значимым агентом передачи гендерных норм и поддержки формирования дифференцированной гендерной идентичности [3].

Обзоры литературы: По мнению Майкла Лэмба, отношения между отцом и ребёнком носят динамический характер и изменяются на протяжении жизненного пути. Эти отношения оказывают как прямое влияние посредством непосредственного взаимодействия, так и косвенное – через функционирование семейной системы в целом. Кроме того, отец играет важную роль в передаче социальных норм и ожиданий, формируя представления ребёнка о допустимом поведении и социальных ролях.

Особое значение имеет качество отцовского участия. Такие характеристики, как эмоциональная теплота, отзывчивость и доступность, рассматриваются в качестве значимых предикторов психологического благополучия ребёнка. Отцы также способствуют успешной адаптации детей, поощряя самостоятельность, исследовательскую активность и устойчивость к новым или сложным жизненным ситуациям.

Влияние отца проявляется и в более широком семейном контексте, где характер отношений между родителями и особенности совместного воспитания определяют степень и характер воздействия отцовского участия на развитие ребёнка. Несмотря на культурные различия в представлениях об отцовстве, многочисленные исследования подтверждают, что отец остаётся важным фактором, оказывающим существенное влияние на различные аспекты развития ребёнка [1].

В работе «Три очерка по теории сексуальности» З. Фрейд вводит концепцию психосексуального развития, рассматривая фаллическую стадию (примерно от трёх до шести лет) как важнейший этап формирования гендерной идентичности. На данном этапе ребёнок осознаёт анатомические различия между полами, что приводит к трансформации эмоциональных привязанностей и формированию основы эдиповой динамики.

Применительно к девочкам Фрейд описывает процесс смещения первоначальной привязанности к матери на фигуру отца, который впоследствии обозначает как «фемининную эдипову установку». Данный переход связан с осознанием половых различий и сопровождается возникновением амбивалентных чувств, включающих привязанность к отцу и соперничество с матерью.

По мнению Фрейда, успешное прохождение данного этапа предполагает разрешение внутреннего конфликта посредством идентификации с родителем своего пола – матерью, что

способствует формированию гендерной идентичности и моральной структуры личности (сверх-Я) [6].

Юнгианский аналитический подход к рассматриваемой проблеме представлен в работах Сьюзан Шварц и Линды Ширс Леонард.

В книге The Absent Father Effect on Daughters Сьюзан Э. Шварц, опираясь на положения аналитической психологии К. Г. Юнга, рассматривает отсутствие отца – как физическое, так и эмоциональное – в качестве одного из ключевых факторов, влияющих на формирование идентичности дочери, её самооценки и моделей межличностных отношений. По мнению автора, отсутствие стабильного отцовского присутствия создаёт глубокий психологический «вакуум», препятствующий формированию устойчивого чувства собственного «Я» и нередко приводящий к неуверенности, фрагментации идентичности и снижению самооценки.

Шварц вводит понятие «жажды отца» («father hunger»), отражающее стремление дочерей к признанию, поддержке и одобрению, которые не были в достаточной мере получены в детстве. Данная потребность оказывает влияние как на сознательные, так и на бессознательные аспекты личностного развития.

Отсутствие отца также нарушает процессы формирования привязанности и способствует возникновению трудностей в построении доверительных, близких и устойчивых отношений. Это, в свою очередь, может приводить к воспроизведению дисфункциональных реляционных паттернов во взрослой жизни. Используя такие юнгианские концепции, как отцовский комплекс и архетипические структуры, Шварц демонстрирует, каким образом интернализированный образ отсутствующего или эмоционально недоступного отца влияет на представления о маскулинности, гендерной идентичности и ожиданиях по отношению к потенциальным партнёрам.

Подобное влияние может проявляться в различных компенсаторных формах поведения, включая идеализацию отсутствующего отца, формирование «ложного Я» или эмоциональное дистанцирование как защитную реакцию на переживание утраты и покинутости.

Кроме того, автор подчёркивает межпоколенческий характер последствий отсутствия отца, указывая на возможность бессознательной передачи неразрешённых отцовских конфликтов следующим поколениям при отсутствии их осознания и психологической проработки. Вместе с тем Шварц отмечает возможность психологического исцеления посредством осознания и интеграции внутреннего образа отца, что способствует формированию более целостной и аутентичной личности. В целом фигура отца, независимо от его фактического присутствия или отсутствия, рассматривается автором как важнейший психологический фактор формирования женской идентичности, моделей отношений и интеграции гендерных аспектов психики [5].

В работе The Wounded Woman Линда Ширс Леонард, также опираясь на юнгианскую традицию, рассматривает отношения между отцом и дочерью как один из ключевых факторов формирования женской идентичности, сексуальности и бессознательных гендерных установок. Отец выступает носителем символической мужской энергии – анимуса, который интернализуется дочерью и впоследствии влияет на её самооценку, уровень автономии и особенности поведения в межличностных отношениях.

По мнению Леонард, позитивные и эмоционально поддерживающие отношения с отцом способствуют формированию устойчивого чувства собственного «Я», уверенности в себе и способности к построению здоровых близких отношений. Напротив, дефицитарные, контролирующие или эмоционально недоступные отцы могут способствовать развитию так называемой «раненой женской идентичности», характеризующейся зависимостью, неуверенностью и искажёнными ожиданиями по отношению к мужчинам.

В подобных случаях интернализированный анимус может приобретать критичный, ригидный или негативный характер, формируя неадаптивные когнитивные и эмоциональные паттерны. Кроме того, автор связывает ранний опыт взаимодействия с отцом с формированием бессознательных сценариев отношений, вследствие чего во взрослом возрасте женщины нередко выбирают партнёров, воспроизводящих неразрешённые динамики отношений с отцом.

Леонард также подчёркивает, что влияние отца на формирование сексуальности дочери носит преимущественно символический характер и осуществляется через систему отношений, формируя представления о маскулинности, собственной ценности и привлекательности. Интеграция феминных и маскулинных аспектов личности рассматривается как необходимое условие психологической зрелости, при этом фигура отца занимает центральное место в данном процессе. В целом Леонард рассматривает отца как значимый психологический фактор процесса индивидуации дочери, определяющий особенности интернализации авторитета, автономии и моделей межличностных отношений [4].

Переходя непосредственно к рассмотрению влияния роли отца на формирование гендерной идентичности ребёнка, следует отметить, что данная проблема получила широкое освещение в трудах многих исследователей.

В сборнике Handbook of Fathers and Child Development подчёркивается значительный вклад отца в различные аспекты развития ребёнка, включая становление гендерной идентичности. Авторы отмечают, что влияние отца осуществляется посредством сложных и многомерных механизмов, формирующих внутренние представления ребёнка о себе и окружающих. Эти представления становятся основой для развития личностной идентичности, в том числе её гендерного компонента. Особая роль отца заключается в предоставлении моделей поведения и способов взаимодействия, отличающихся от материнских, что способствует усвоению ребёнком гендерных ролей и социальных ожиданий.

Кроме того, присутствие отца и степень его вовлечённости оказывают существенное влияние на эмоциональную и социальную регуляцию ребёнка, которые тесно связаны с процессами формирования идентичности и ролевой ориентации [3].

Схожую позицию занимает Генри Б. Биллер, который рассматривает отца как центральную фигуру в процессе формирования гендерной идентичности ребёнка посредством механизмов идентификации, моделирования и подкрепления. В его теоретической концепции отец выступает в качестве основной референтной фигуры, особенно для мальчиков, способствуя усвоению мужских ролей, моделей поведения и системы ценностей. Вместе с тем влияние отца распространяется и на дочерей, формируя их представления о маскулинности и ожидания относительно взаимоотношений с мужчинами.

По мнению Биллера, присутствие вовлечённого и эмоционально отзывчивого отца способствует развитию устойчивой и дифференцированной гендерной идентичности, тогда как его отсутствие, непоследовательность или эмоциональная дистанцированность могут приводить к неопределённости в освоении гендерных ролей. Отец также закрепляет соответствующие гендеру модели поведения посредством ожиданий, обратной связи и дифференцированного отношения, оказывая влияние на формирование полоролевой ориентации ребёнка.

Более того, характер отцовского участия во многом определяет особенности гендерно-ролевых установок детей. Поддерживающее и эмоционально сбалансированное отцовство связано с формированием более гибких и адаптивных представлений о гендерных ролях. В целом Биллер рассматривает отца как один из ключевых факторов когнитивного, эмоционального и социального развития, посредством которого ребёнок формирует представление о гендере и собственном месте в системе гендерных отношений [2].

Особый интерес к проблеме формирования гендерной идентичности проявлял и Зигмунд Фрейд. В работе «Женская сексуальность» он развивает представления о психосексуальном развитии девочек и подробно анализирует особенности женской версии эдипова комплекса. Фрейд утверждает, что первоначально девочка испытывает первичную привязанность к матери, однако в дальнейшем её либидозное внимание смещается на фигуру отца. Данный переход рассматривается как важнейшее структурное изменение в развитии девочки и связывается с осознанием анатомических различий между полами, которое в психоаналитической теории обозначается понятием «зависть к пенису».

В отличие от мальчиков, девочки, по мнению Фрейда, переживают эдипов конфликт иначе: его разрешение происходит менее определённо, а процесс идентификации с матерью отличается большей сложностью и продолжительностью. Учёный приходит к выводу, что успешное разрешение данной динамики способствует формированию фемининной идентичности, развитию моделей межличностных отношений и становлению сверх-Я, хотя структура и последствия этого процесса существенно отличаются от мужского варианта развития [7].

Таким образом, анализ представленных теоретических подходов позволяет сделать вывод о значимой роли отца в формировании гендерной идентичности ребёнка. Для мальчиков отец выступает образцом мужского поведения, уверенности и ответственности, посредством которого они осваивают способы эмоционального самовыражения и взаимодействия с противоположным полом. Для девочек отец становится первым опытом взаимодействия с мужчиной, формируя представления о мужском поведении и ожидания относительно будущих отношений.

Отсутствие отца либо его эмоциональная холодность могут затруднять формирование полоролевого поведения и негативно сказываться на самооценке ребёнка. В свою очередь, гиперкритичность или агрессивность со стороны отца способны вызывать чувство страха и эмоциональной отстранённости, оказывая влияние на характер будущих межличностных отношений. Напротив, в семьях, где отец демонстрирует уважение, поддержку и эмоциональную вовлечённость, дети чаще формируют устойчивую гендерную идентичность и позитивное отношение к собственной половой роли. Активное участие отца в воспитании способствует более гармоничному освоению гендерных ролей и снижает вероятность формирования стереотипных искажений.

Список литературы:

1. The role of the father in child development / edited by Michael E. Lamb. — 4th ed., 2004
2. Father, Child and Sex Role: Paternal Determinants in Personality Development (Lexington Books) / Biller, Henry B., 1971
3. Handbook of Fathers and Child Development: Prenatal to Preschool / Hiram E. Fitzgerald, Kai von Klitzing, Natasha J. Cabrera, Júlia Scarano de Mendonça, Thomas Skjøthaug, 2020
4. The Wounded Woman, Healing the Father-Daughter Relationship / Linda Schierse Leonard, 1982
5. The Absent Father Effect on Daughters / Susan E. Schwartz, 2020
6. Three Essays on the Theory of Sexuality / Sigmund Freud, 1905
7. Female Sexuality / Sigmund Freud, 1931

Economic Sciences

Social Media Marketing Practices of Kazakhstan's Local Coffee and Tea Chains: A Multiple-Case Analysis Based on the S-O-R Model

WEI Mengyun
Farabi International Business School, Al-Farabi Kazakh National University, Almaty, Kazakhstan

Abstract

Purpose. This paper investigates how three leading local coffee and tea chains in Kazakhstan design and deploy social media marketing (SMM) stimuli, and how the observed patterns can be interpreted through the Stimulus-Organism-Response (S-O-R) theoretical framework. Systematic academic analysis of local Central-Asian F&B chains remains scarce, and the study responds to this gap.

Design/Methodology. The paper adopts a multiple-case study design following Yin (2018). Three purposively selected Kazakhstani chains — Coffee BOOM, Marrone Rosso and Daily Coffee — jointly cover the strategic space of the local coffee/tea market along the value-to-premium spectrum. Data were collected from the brands' public Instagram, TikTok and Telegram accounts, supplemented by industry press and the DataReportal Digital 2025 Kazakhstan report. Cross-case pattern-matching was used to map each brand's SMM configuration onto the S-O-R model's stimulus, organism and response stages.

Findings. The three brands adopt visibly distinct SMM configurations that map to different S-O-R conversion pathways. Coffee BOOM emphasises reach-oriented stimuli (breakfast promotions, mass-audience Instagram posts) that convert primarily through perceived value. Marrone Rosso relies on aesthetic-quality stimuli and Instagram lifestyle photography, converting mainly through brand attitude. Daily Coffee operates a youth-culture, atmosphere-driven pattern whose primary conversion route runs through affective brand attachment — a finding that departs from the dominant logic of the Chinese SMM literature. Three sector-level gaps are identified: under-utilisation of TikTok despite its dominance in Kazakhstan's social media landscape; fragmented private-domain retention via Telegram; and limited structured cooperation with local micro-influencers (KOCs).

Originality/Value. This is, to the author's knowledge, the first study to apply the integrated S-O-R and AISAS framework to Kazakhstan's local F&B sector. The findings offer an empirical baseline against which incoming international competitors — most visibly Starbucks, Costa Coffee, and Chinese chains such as Mixue Bingcheng — can be benchmarked, and they identify three concrete, actionable levers through which local brands can defend and grow market share.

Keywords. social media marketing; S-O-R model; coffee and tea chains; Kazakhstan; multiple-case study; consumer behaviour.

1. Introduction

The Kazakhstani coffee and tea market has evolved rapidly. What began in the late 1990s as a small, largely imported category has become one of the most dynamic segments of the country's food-and-beverage retail landscape. Industry observers report that the number of coffee-selling outlets in Almaty increased by more than 60 per cent between 2019 and 2022, as domestic consumer preferences shifted from at-home tea consumption toward out-of-home coffee culture [1]. Digital infrastructure has expanded on a parallel trajectory. By early 2025, Kazakhstan had 19.2 million internet users (92.9 per cent penetration) and 15.7 million active social-media users (75.7 per cent penetration), with TikTok, Instagram and Telegram emerging as the three dominant consumer platforms [2]. For local coffee and tea chains, the practical consequence of this dual transformation is straightforward: the marginal customer is now increasingly won or lost through the social-media touchpoint rather than through the physical retail one.

The domestic competitive environment has become materially more contested. International chains — Costa Coffee (since 2012), Starbucks (since 2015), and, more recently, Chinese and Brazilian entrants such as Mixue Bingcheng and The Coffee — have expanded their Kazakhstani footprints, bringing with them mature SMM playbooks refined in larger and more sophisticated home markets. Local Kazakhstani chains now face the strategic question of how to defend and grow share in a market where the international competitors' SMM capabilities are, at least initially, stronger than their own. Answering that question requires a first step: a systematic characterisation of what local chains are actually doing on social media.

The research question of this study is stated as follows. How do leading Kazakhstani coffee and tea chains configure their social media marketing stimuli, and how do the observed configurations map onto the stimulus, organism and response stages of the S-O-R theoretical model? The paper contributes to the literature in three ways. It provides the first published multiple-case analysis of SMM practices among Kazakhstani local F&B chains, applies the integrated S-O-R and AISAS framework — well-established in the Chinese consumer marketing context [3] — to a new geographic setting, and identifies three sector-level SMM gaps that hold direct managerial relevance for local brands preparing to defend against incoming international competition.

The remainder of the paper is organised as follows. Section 2 sets out the theoretical framework, integrating the S-O-R and AISAS models. Section 3 describes Kazakhstan's digital marketing environment and the local coffee/tea market. Section 4 explains the multiple-case research design, case selection and data sources. Section 5 presents the three case analyses individually and then cross-case comparatively. Section 6 discusses the three sector-level findings. Section 7 offers managerial implications, and Section 8 concludes with limitations and directions for future research.

2. Theoretical Framework

2.1. Social media marketing in the food and beverage sector

Social media marketing refers to the deliberate use of social-networking sites and short-video platforms to build brand-consumer relationships, disseminate branded content and drive commercial outcomes [4]. Within the F&B sector, SMM has become a dominant marketing paradigm for two structural reasons. F&B consumption is inherently social — a coffee, a cup of tea, a meal — and therefore lends itself naturally to the visual, shareable content genres that social platforms reward. Furthermore, the low unit price of F&B products makes the marginal cost per media impression a critical determinant of campaign viability, which favours the low-cost content genres that ordinary consumers themselves produce and distribute [5].

The academic literature on F&B SMM has accumulated substantial evidence on the effects of content characteristics, platform choice and influencer credibility on consumer response. De

Vries, Gensler and Leeflang (2012) demonstrated that vivid, interactive brand posts on fan pages generate materially higher engagement than plain informational posts [6]. Berger and Milkman (2012) established that content evoking high-arousal emotions — awe, joy, anger — is more likely to be shared than content evoking low-arousal emotions [7]. Kim and Ko (2012), in an influential study of luxury fashion SMM, isolated five constructs of perceived SMM activity — entertainment, interaction, trendiness, customisation and word of mouth — that jointly enhance customer equity and purchase intention [8]. Hollebeek, Glynn and Brodie (2014) developed the consumer brand engagement (CBE) scale, which has since become the standard operationalisation of interactive brand-consumer relationships in social media settings [9]. More recently, Djafarova and Rushworth (2017) documented that ordinary online micro-celebrities exert stronger purchase-decision influence on young consumers than traditional celebrity endorsers, on account of perceived authenticity and credibility [10].

2.2. The Stimulus-Organism-Response (S-O-R) framework

The S-O-R model, originally developed by Mehrabian and Russell (1974) within environmental psychology, posits that external stimuli (S) act on an individual's internal cognitive-affective state (O) to produce behavioural responses (R) [11]. Donovan and Rossiter (1982) subsequently adapted the framework to retail environments, providing the foundation for its later application to digital marketing contexts [12]. In an SMM setting, the three stages of the model are typically operationalised as follows. The stimulus stage comprises the marketing content, influencer endorsements and platform-interaction opportunities the brand deploys. The organism stage comprises the consumer's evolving brand attitude and perceived value. The response stage comprises the resulting purchase intention and actual purchase behaviour.

The S-O-R framework suits the F&B SMM context well because it accommodates both affective (brand attitude) and cognitive (perceived value) routes to purchase. This matters because different F&B categories — mass value products, premium quality products, cultural or lifestyle products — activate these two routes in different proportions.

2.3. The AISAS model as a process-stage complement

While the S-O-R model captures the causal architecture of marketing influence, the AISAS model — proposed by the Japanese advertising agency Dentsu — captures the sequential stages through which a consumer moves in a digitally-mediated purchase decision: Attention, Interest, Search, Action, and Share [14]. AISAS replaces the desire and memory stages of the older AIDMA framework with search and share, reflecting the reality that in the internet era consumers actively investigate brands before purchase and actively communicate about them afterward. Integrating AISAS with the S-O-R model allows the analysis of SMM practices to trace which stage of the consumer journey a given brand's stimulus configuration is optimised for. Attention-stage stimuli (mass reach content) look very different from search-stage stimuli (in-depth lifestyle notes) or share-stage stimuli (user-generated content invitations).

2.4. Integrated analytical framework

The present paper adopts the integrated S-O-R plus AISAS framework as its analytical lens. SMM stimuli are classified along three primary dimensions — content creativity, influencer credibility and platform interactivity — that jointly determine the strength of the stimulus signal reaching the consumer. These stimuli act on two organism constructs, brand attitude and perceived value, whose relative weight depends on the brand's positioning. The AISAS stages, superimposed on this architecture, indicate which point in the consumer journey each stimulus configuration is designed to activate.

Digital transactional interfaces — mobile ordering apps, delivery platforms, mini-programmes — clearly matter for the organism-to-response transition, and the Technology Acceptance Model (TAM) tradition, in which perceived ease of use is a well-established construct [13], provides useful vocabulary for describing this transition. This paper acknowledges the

interface layer but treats it as contextual rather than as a formally moderating variable, because the empirical material analysed here does not permit meaningful evaluation of ease-of-use effects at the consumer level. This framework, so specified, is applied in Section 5 as the coding scheme for the multiple-case analysis.

3. Kazakhstan's Digital Marketing Environment

3.1. Digital penetration and platform landscape

By early 2025, Kazakhstan reported one of the highest levels of digital penetration in Central Asia. The DataReportal Digital 2025 Kazakhstan report documents 19.2 million internet users (92.9 per cent of the total population), 15.7 million active social media user identities (75.7 per cent of the population), and 26.6 million mobile connections (128 per cent of the population, indicating widespread multi-device or dual-SIM use) [2]. Average mobile internet speed rose 51 per cent year-on-year to 53.7 Mbps, and the country's total advertising market grew 17 per cent in the first half of 2025 to KZT 27 billion, with digital advertising alone contributing approximately KZT 14 billion, up 29 per cent year-on-year [15].

The platform landscape has three defining features. TikTok has become the largest social platform by advertising reach, with 15.7 million adult users at the start of 2025 — a reach approximately equal to the entire adult population, reflecting both very high engagement and multi-account use [2]. Instagram remains the country's most influential visual platform. It had 12.4 million adult users in early 2025 rising to 13.1 million (63 per cent of the population) by late 2025, and more than 90 per cent of Kazakhstani internet users aged 18 and above visit Instagram monthly [16]. Telegram, meanwhile, has evolved from a messenger into a full commercial ecosystem: channels, bots and mini-applications increasingly function as brands' private-domain retention infrastructure, and the platform is now the primary news and information source for a substantial share of Kazakhstani young adults [17]. Facebook, by contrast, remains a comparatively small platform in Kazakhstan (2.6 million users, unchanged year-on-year), and YouTube, though widely used for entertainment consumption, plays a smaller role in F&B brand marketing than the three platforms above.

3.2. The local coffee and tea market

Kazakhstan is traditionally a tea-drinking country — black tea alone accounts for approximately 84 per cent of the domestic tea market — but out-of-home coffee consumption has grown rapidly since the mid-2010s [1]. The market entry of international coffee chains (Costa Coffee in 2012, Starbucks in 2015) accelerated a shift in urban consumer preferences, and by 2022 the number of coffee-selling points in Almaty had risen by approximately 120 per cent relative to 2019 [1]. Local Kazakhstani chains have expanded in parallel, occupying strategic positions ranging from mass-market breakfast-and-coffee formats to premium Italian-inspired lifestyle cafes and boutique youth-oriented spaces. The competitive environment tightened further in 2024–2026 with the arrival of Chinese chains — Mixue Bingcheng entered the Central Asian market, and Brazilian premium chain The Coffee opened its first Central Asian outlet in Almaty in January 2026 [18].

3.3. Regulatory and infrastructural context

Two contextual features of the Kazakhstani market shape the SMM strategies available to local F&B brands. Kazakhstan lacks a domestic super-app equivalent to China's WeChat, so brand private-domain traffic operations must be built on multi-platform combinations — typically Instagram for discovery, TikTok for reach, and Telegram for retention — rather than concentrated within a single platform ecosystem. In addition, Kaspi.kz, the country's dominant mobile-payment and digital-services platform, has integrated ordering and payment infrastructure that some local chains have begun to leverage as a transactional back-end, effectively serving the role that WeChat mini-programmes serve in the Chinese context. The implications of these structural features for SMM practice are elaborated in the case analyses in Section 5.

4. Research Method

4.1. Research design

The study adopts a multiple-case study design following the methodology set out by Yin (2018) [19]. A multiple-case design is appropriate here for three reasons. The research question is exploratory rather than confirmatory — the objective is to characterise how local Kazakhstani chains configure their SMM rather than to test a pre-specified hypothesis. The analytical unit is the brand-level marketing configuration, best captured through in-depth examination of a small number of information-rich cases rather than through large-sample statistical inference. Finally, the multiple-case design permits cross-case pattern matching, which is the analytical technique through which the study identifies sector-level regularities in Section 6.

4.2. Case selection

Three Kazakhstani coffee and tea chains were selected through theoretical sampling on the basis of three criteria: (i) each brand is Kazakhstani-founded and operates primarily within the domestic market; (ii) each maintains an active public presence on at least Instagram, with observable content activity within the twelve months preceding the study; (iii) the three brands together span the strategic space of the local market along the price-and-positioning spectrum, from mass-value to premium-lifestyle. Table 1 summarises the three selected cases.

Table 1 – Case selection: three Kazakhstani coffee and tea chains

Brand	Positioning	Format	Primary social channel
Coffee BOOM	Mass-market, value-driven	Breakfast and Western-style fast food + coffee	Instagram
Marrone Rosso	Mid-to-high-end, Italian-inspired	European-style cafe: coffee, pastries, salads	Instagram
Daily Coffee	Boutique, youth-culture	Single-location trendy cafe with large terrace	Instagram

Source: compiled by the author based on brand-level public information and industry press [1, 20].

The three brands together satisfy Yin's minimum-cases requirement (three cases for replication logic) and cover the strategic space of Kazakhstan's local coffee sector along both price positioning and marketing paradigm. Coffee BOOM anchors the mass end with a nationwide multi-city footprint; Marrone Rosso occupies the mid-to-high-end aesthetic-quality position with locations in Almaty, Astana and Shymkent; Daily Coffee represents the boutique youth-culture segment. The strategic-space logic parallels published multiple-case studies of Chinese consumer marketing in the digital era [3], enabling comparative interpretation between the Kazakhstani findings and the Chinese benchmark.

4.3. Data sources

In line with Yin's multiple-sources-of-evidence principle, data were collected from three distinct source types to allow triangulation. First, brand-produced public content: each brand's official Instagram, TikTok (where present) and Telegram (where present) accounts were reviewed for content volume, format mix, posting frequency, and engagement patterns. Second, industry and business press: articles from Astana Times, Kursiv, Tengri News, STiR Coffee and Tea Magazine, and DKNews were used to establish brand history, store network scale and market context. Third, macro-level digital environment data: the DataReportal Digital 2025 Kazakhstan report and secondary industry analyses were used to establish the platform-level context within

which brand-level activity takes place [2, 15, 16]. The field observation window spans late 2025 through early 2026.

4.4. Data analysis

Data analysis proceeded in two stages. In the first stage, each case was analysed individually using the integrated S-O-R and AISAS framework as the coding scheme. For each brand, evidence was assembled on the three stimulus dimensions (content creativity, influencer credibility, platform interactivity), the two organism constructs on which the stimuli appear to act (brand attitude, perceived value), and the response-stage transactional interface. In the second stage, cross-case pattern matching was applied to identify commonalities and divergences across the three brands. The analytical output of this two-stage process is presented in Section 5. The study's internal validity is supported by data-source triangulation; external validity, in the multiple-case study logic, rests on analytical rather than statistical generalisation to the theoretical framework rather than to the Kazakhstani market as a whole.

5. Case Analysis

5.1. Coffee BOOM: mass-market value-driven pattern

Coffee BOOM is a well-established Kazakhstani cafe chain with a nationwide footprint that includes multiple locations in Almaty, Astana and other regional cities. Its brand proposition combines coffee with breakfast and Western-style fast food (burgers, avocado toast, pastries), positioning the chain in the mass-market segment where the primary consumer promise is a convenient, comfortable urban meal environment at a moderate price point [20]. On Instagram, the brand's official Almaty account (@coffeeboom.almaty) publishes primarily food-and-beverage product photography, seasonal breakfast promotions and store-atmosphere posts. Content aims to communicate value-for-money and everyday accessibility rather than aspirational lifestyle.

Read through the S-O-R lens, Coffee BOOM's SMM logic can be characterised as follows. The stimulus configuration emphasises perceived-value content — price-and-portion promotions, breakfast deals, product bundles — over aesthetic-quality content or influencer endorsements. The organism-stage effect operates primarily through the perceived-value route rather than the brand-attitude route. Consumers are recruited on the strength of a value proposition (a substantial breakfast for a reasonable price at a convenient location) rather than on aspirational brand identification. The response-stage conversion is straightforward and immediate — the walk-in visit — with limited investment in digital-transactional infrastructure or private-domain retention. In AISAS terms, the brand's SMM output is optimised for the Attention and Action stages of the consumer journey, with under-developed activity in the Interest, Search and Share stages.

The pattern has clear strengths: the value proposition is legible and the operational model is scalable, which has enabled the nationwide expansion. Its weaknesses, viewed through the analytical framework, are also clear. The under-investment in the Search and Share stages means Coffee BOOM captures relatively little of the discovery traffic generated by Kazakhstani consumers searching Instagram or TikTok for cafe recommendations. Furthermore, the absence of a structured private-domain retention layer means repeat-purchase mechanics rely primarily on location convenience and habitual behaviour rather than on brand-owned communication channels.

5.2. Marrone Rosso: mid-to-high-end aesthetic-quality pattern

Marrone Rosso is an Italian-inspired cafe chain operating in Almaty, Astana and Shymkent, occupying the mid-to-high-end segment of the Kazakhstani coffee market. Founded in 2009, the brand's positioning centres on European cafe atmosphere — the name itself references Italian visual culture — with a menu that combines specialty coffee, Italian-style pastries, sandwiches and salads, served in cafe environments designed for both casual meetings and remote work [21]. The brand's Instagram output (@marrone_rosso, approximately 22,000 followers as of early 2026, with more than 2,300 archived posts) emphasises visual product presentation (coffee art, plated

pastries), cafe interior aesthetics, and lifestyle association imagery. Content quality and photography standards are notably higher than those observable at mass-market competitors, and the brand's established seventeen-year presence gives its Instagram feed a coherent long-arc visual identity.

Marrone Rosso's SMM pathway, in the S-O-R framework, operates primarily through the brand-attitude route rather than the perceived-value route. The stimulus configuration emphasises aesthetic-quality content — high-production-value product photography, seasonal menu launches with strong visual identity, cafe-interior lifestyle imagery — which is designed to build a positive affective brand association rather than to communicate an immediate price advantage. The organism-stage effect concentrates on brand attitude. A Marrone Rosso consumer is expected to develop a favourable identity association with the brand's European-cafe aesthetic, which then justifies the premium price relative to mass-market alternatives. In AISAS terms, the brand is strongest at the Interest and Search stages — a consumer prompted by an Instagram post can discover the brand's visual identity in a coherent, curated feed — but weaker at the Attention stage (limited high-frequency reach content) and at the Share stage (limited structured user-generated-content invitations).

This pattern aligns closely with the SMM configuration that Kim and Ko (2012) documented for luxury fashion brands: entertainment and trendiness content combined with lifestyle association [8]. What the Kazakhstani F&B context adds is a boundary condition. Because the Kazakhstani coffee market is smaller and more competitively contested than luxury fashion, the premium-attitude mechanism can only sustain its price premium if aesthetic quality is maintained consistently across every touchpoint. Occasional inconsistencies visible in consumer review sources — variable pastry freshness, inconsistent service quality across locations — represent the specific vulnerability of the aesthetic-quality pathway, because a single negative brand-attitude signal can undermine substantially more customer-equity value than an equivalent value-stage disappointment would.

5.3. Daily Coffee: boutique youth-culture atmospheric pattern

Daily Coffee occupies a distinct third strategic position. It is a boutique, atmospherically-driven youth cafe located centrally in Almaty at the intersection of Abylai Khan Avenue and Abai Avenue, with a large terrace that has become the brand's signature physical asset [22]. The brand does not attempt nationwide expansion; its competitive strategy is built on being a specific place — a scene — rather than a scalable format. Its Instagram account (@dailycoffee.almaty) reflects this positioning, with content emphasising the terrace atmosphere, the young clientele, the surrounding urban environment, and the fusion of coffee, food and cocktail offerings served across long opening hours.

In the S-O-R framework, Daily Coffee's mechanism operates through a third organism-stage route that neither of the two preceding brands emphasises: affective brand attachment rooted in place-and-scene identification. The stimulus configuration is atmospheric rather than either value-oriented or aesthetically-precious. Content shows people in the space, the space in various weather and light conditions, the space at different times of day. The organism-stage effect combines brand attitude and perceived value in an integrated way. Consumers pay for the experience of being at Daily Coffee, which is neither a pure value proposition nor a pure aesthetic proposition but a complex bundle of atmosphere, sociability and youth-cultural identification. The response-stage conversion is spatially localised. Daily Coffee does not aim to convert consumers to repeat purchase through a delivery app or a membership programme, but to convert them into recurring visitors to the specific physical location.

In AISAS terms, Daily Coffee is unusually strong at the Share stage: user-generated content of the terrace circulates organically on Instagram, functioning as the brand's most valuable and most authentic marketing output. This mechanism relies on a specific consumer segment —

Almaty's Gen-Z and young-millennial urban demographic — and is not easily portable to a nationwide expansion model.

The Daily Coffee case is, in the researcher's reading, the most theoretically interesting of the three. Studies of Chinese consumer marketing, on which the integrated S-O-R and AISAS framework has been most extensively applied, tend to treat brand attitude and perceived value as the two organism-stage constructs that jointly mediate stimulus-to-response conversion [3]. Daily Coffee sits awkwardly on that map. Its conversion does not proceed through a value calculation, nor through the kind of aspirational brand identification that premium chains typically cultivate. Instead, the object of consumer attachment is neither the product nor the brand identity in the abstract, but the physical scene of the terrace itself, with the coffee acting almost as a ticket of admission. This scene-and-atmosphere pathway, though not entirely absent from the Chinese literature, is treated there as a secondary or supporting mechanism rather than as a primary route. In the Kazakhstani setting it appears to function as a first-order competitive strategy, and this asymmetry warrants attention in later work.

5.4. Cross-case comparison

Table 2 sets the three brands' SMM mechanisms side by side across the analytical dimensions of the integrated framework. Read across, the three brands do not converge on a single dominant SMM configuration but distribute themselves across three distinct combinations of stimulus type, organism-stage route and AISAS-stage emphasis. This distribution is the central analytical finding of the case analysis: the strategic space of the Kazakhstani local coffee market is heterogeneous, and each of the three studied brands occupies a coherent internally-consistent position within that space.

Table 2 – Cross-case comparison of the three brands' SMM mechanisms

Dimension	Coffee BOOM	Marrone Rosso	Daily Coffee
Positioning	Mass-market value	Mid-high aesthetic	Boutique youth-culture
Primary stimulus type	Value/promotion content	Aesthetic/lifestyle content	Atmospheric/scene content
Dominant organism route	Perceived value	Brand attitude	Affective attachment (blended)
Strong AISAS stage(s)	Attention, Action	Interest, Search	Share
Weak AISAS stage(s)	Search, Share	Attention, Share	Search
Instagram content emphasis	Product+price posts	Curated product photography	Terrace+atmosphere posts
TikTok activity	Limited	Limited	Limited
Telegram private-domain	Under-developed	Under-developed	Under-developed
KOL/KOC cooperation	Occasional	Occasional	Organic UGC-driven

Source: compiled by the author based on public content review of each brand's official accounts, cross-referenced with industry press and consumer-review sources [20, 21, 22].

Two cross-case observations sharpen the strategic reading of Table 2. Notwithstanding the three brands' distinct positioning strategies, all three under-invest in the same three areas: structured TikTok content production, Telegram-based private-domain retention, and structured cooperation with local micro-influencers. These common under-investment areas constitute the sector-level SMM gap discussed in Section 6. In addition, none of the three brands has adopted the integrated public-to-private-domain SMM architecture that has become the dominant model among leading Chinese consumer brands [3] — one in which public-platform stimuli feed directly into a private-domain retention layer within which repeat-purchase mechanics are operated. The absence of this architecture is partly attributable to Kazakhstan's lack of a WeChat-equivalent super-app, but the Telegram ecosystem provides at least a partial functional substitute that none of the three brands is presently exploiting at scale.

6. Discussion

The cross-case analysis identifies three sector-level SMM gaps that hold across the three studied brands despite their distinct strategic positions. Each gap is discussed in turn, with reference to the theoretical framework and to comparable findings from the international SMM literature.

6.1. The TikTok gap

TikTok is by advertising reach the largest social platform in Kazakhstan, with 15.7 million adult users at the start of 2025 [2]. Yet across the three studied brands, TikTok activity is limited — none of the three operates a high-frequency, professionally-produced TikTok content channel comparable to their Instagram activity. This under-utilisation is theoretically consistent with the AISAS-stage emphasis mapped in Table 2. TikTok's algorithmic short-video feed is optimised for the Attention stage, and only Coffee BOOM among the three brands operates a logic whose primary strength is at the Attention stage. That said, the observed pattern is an under-response to the Kazakhstani platform reality. With TikTok reaching essentially the entire adult population, absence from the platform means voluntary forfeiture of the largest available discovery surface. The finding parallels the pattern documented in the Chinese consumer marketing literature, where mass-market brands derive most of their top-of-funnel reach from short-video platforms such as Douyin (TikTok's Chinese-market equivalent) [3]. Kazakhstani local brands have not yet developed the equivalent TikTok-native content strategy.

6.2. The Telegram private-domain gap

Telegram has evolved in the Kazakhstani market from a messenger into a full commercial ecosystem, with channels, bots and mini-applications increasingly serving as brands' private-domain retention infrastructure [17]. The theoretical rationale for private-domain traffic operation, established in the Chinese SMM literature, is that a brand that has spent scarce campaign budget to convert a consumer through a public-platform stimulus has a strong incentive to retain that consumer in a channel it owns and controls, rather than to leave the retention step to the algorithm of the platform on which the stimulus was received [3, 9]. Across the three studied brands, this private-domain layer is under-developed. Consumer retention appears to rely primarily on location convenience (Coffee BOOM), brand-attitude loyalty (Marrone Rosso) or scene-specific attachment (Daily Coffee), rather than on brand-owned communication channels through which promotional stimuli can be delivered directly to identified past customers. Given the maturity of the Telegram ecosystem in Kazakhstan, this gap represents a substantial unexploited SMM opportunity for local brands.

6.3. The KOC cooperation gap

Djafarova and Rushworth (2017) documented that ordinary online micro-celebrities exert stronger purchase-decision influence on young consumers than traditional celebrity endorsers, on account of perceived authenticity and credibility [10]. In the Chinese SMM context, this insight has been operationalised at scale through structured KOC (key opinion consumer) cooperation programmes — brands contract simultaneously with dozens of mid- and long-tail influencers to generate high-volume, credible-feeling brand content [3]. Across the three studied Kazakhstani brands, structured KOC cooperation of this kind is largely absent. Influencer content that mentions the three brands appears predominantly to be organic user-generated content, not commissioned KOC content. This gap is theoretically the most consequential of the three, as it lies at the intersection of the stimulus-side credibility mechanism and the KOC-cost economics — the per-impression cost of KOC-generated content is materially lower than the per-impression cost of professional advertising.

7. Managerial Implications

The three sector-level SMM gaps identified in Section 6 translate directly into three managerial recommendations for Kazakhstani local coffee and tea chains. The recommendations are stated at the sector level and can be operationalised differently by brands occupying different strategic positions.

First, Kazakhstani local brands should invest in TikTok content channels calibrated to their strategic position. For a mass-market brand such as Coffee BOOM, this implies short-video content aimed at Attention-stage reach — 15- to 30-second product-and-atmosphere clips using platform-

native trends and music. For premium-positioned brands such as Marrone Rosso, TikTok can be used more selectively to showcase seasonal menu launches or store-atmosphere content, complementing the primary Instagram feed. The under-investment in TikTok that this study documents represents an addressable competitive weakness, given that incoming international competitors — most visibly Chinese chains such as Mixue Bingcheng — are known to prioritise short-video reach at market entry.

Furthermore, Kazakhstani local brands should build Telegram-based private-domain retention infrastructure. At a minimum, this means operating an official Telegram channel that publishes brand updates, promotional codes and event announcements. More ambitiously, it means integrating a Telegram bot with the brand's ordering and loyalty systems, enabling coupon distribution and repeat-purchase mechanics to be delivered directly to identified past customers. Because Telegram is already the primary information platform for a substantial share of Kazakhstani young adults, the marginal cost of building a brand's initial private-domain audience through this channel is comparatively low, and the retention economics improve materially once the mechanism is operational.

In addition, Kazakhstani local brands should develop structured cooperation frameworks with local Instagram and TikTok micro-influencers. The structured KOC playbook developed in the Chinese context — batches of ten-to-twenty mid- and long-tail creators contracted with brief content templates and moderate per-post fees — is directly portable to the Kazakhstani market. The primary adaptation is the substitution of Instagram and TikTok for the Chinese Xiaohongshu and Douyin. The cost economics are particularly favourable for local brands whose per-outlet marketing budget is limited: a KOC roster of ten to fifteen Almaty-based creators can generate materially higher content volume and credibility-weighted reach than a comparable investment in traditional advertising.

8. Conclusion, Limitations and Future Research

This paper has applied the integrated S-O-R and AISAS framework to a multiple-case analysis of three leading Kazakhstani local coffee and tea chains. The analysis produces three main contributions. It characterises three distinct SMM pathways operating in the Kazakhstani local coffee market — a mass-market value pattern (Coffee BOOM), an aesthetic-quality pattern (Marrone Rosso), and a boutique atmospheric pattern (Daily Coffee) — each internally consistent and each addressing a different consumer segment. It identifies three sector-level SMM gaps common to all three brands: under-investment in TikTok, under-investment in Telegram-based private-domain retention, and absence of structured KOC cooperation. Furthermore, it derives three concrete managerial recommendations that follow directly from the identified gaps and that provide an actionable defensive playbook against incoming international competition.

The study has three main limitations. The multiple-case design is inherently qualitative and does not permit statistical generalisation to the Kazakhstani coffee market as a whole; extrapolation of the findings to other local chains rests on analytical rather than statistical logic. In addition, the data sources are limited to publicly observable brand content and secondary industry sources. A follow-up study incorporating primary consumer survey data would materially strengthen the response-stage evidence base and permit direct testing of the underlying S-O-R hypothesis system. Nevertheless, the qualitative baseline established here is a necessary precondition for such quantitative work. The study also captures the SMM landscape at a specific point in time (late 2025 through early 2026). Given the pace of change in Kazakhstan's digital ecosystem — the arrival of Chinese and Brazilian chains, the growth of TikTok, the evolution of the Telegram commercial ecosystem — the empirical findings will require periodic updating.

Three directions for future research follow from these limitations. A follow-up consumer-survey study operationalising the S-O-R hypothesis system with a Kazakhstani Gen-Z sample would allow direct empirical testing of which stimulus dimensions most strongly drive purchase intention

in the local market. In addition, a comparative study of local Kazakhstani chains versus incoming international entrants — specifically the newly-arrived Chinese and Brazilian chains — would allow the SMM playbooks of the two brand groups to be compared side by side, with implications for both academic understanding of SMM cross-market portability and practical guidance for local defenders. Furthermore, the Daily Coffee finding — that a scene-and-atmosphere pathway can function as a first-order competitive strategy in a way not centrally documented in the Chinese literature — invites dedicated theoretical work on how physical-place attachment interacts with digital-content strategy in emerging Central-Asian consumer markets.

References

1. STiR Coffee and Tea Magazine. (2022). In tea-loving Kazakhstan, a coffee culture grows. Retrieved from https://stir-tea-coffee.com
2. DataReportal, We Are Social, & Meltwater. (2025). Digital 2025: Kazakhstan. Retrieved from https://datareportal.com/reports/digital-2025-kazakhstan
3. Sohaib, M., Safeer, A. A., & Majeed, A. (2022). Role of social media marketing activities in China's e-commerce industry: A stimulus organism response theory context. Frontiers in Psychology, 13, 941058.
4. Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. Business Horizons, 53(1), 59–68.
5. Kotler, P., & Keller, K. L. (2016). Marketing management (15th ed.). Harlow: Pearson Education.
6. De Vries, L., Gensler, S., & Leeflang, P. S. H. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. Journal of Interactive Marketing, 26(2), 83–91.
7. Berger, J., & Milkman, K. L. (2012). What makes online content viral? Journal of Marketing Research, 49(2), 192–205.
8. Kim, A. J., & Ko, E. (2012). Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. Journal of Business Research, 65(10), 1480–1486.
9. Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer brand engagement in social media: Conceptualization, scale development and validation. Journal of Interactive Marketing, 28(2), 149–165.
10. Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. Computers in Human Behavior, 68, 1–7.
11. Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology. Cambridge, MA: MIT Press.
12. Donovan, R. J., & Rossiter, J. R. (1982). Store atmosphere: An environmental psychology approach. Journal of Retailing, 58(1), 34–57.
13. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.
14. Dentsu Inc. (2005). AISAS: A new consumer behaviour model for the cross-media era. Tokyo: Dentsu Inc.
15. BYYD. (2025). What awaits digital marketing in Kazakhstan in 2025. Retrieved from https://www.byyd.me/en/blog/
16. The Astana Times. (2026, January). How social media is shaping Kazakhstan's new generation. Retrieved from https://astanatimes.com
17. Silk Road Studies Program. (2025). Youth and digital technology in Central Asia. Retrieved from https://www.silkroadstudies.org
18. DKNews. (2026, January). The Coffee enters Kazakhstan: Global brand opens its first Central Asia café in Almaty. Retrieved from https://dknews.kz

19. Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). Thousand Oaks, CA: Sage Publications.
20. Coffee BOOM. (2025). Official Instagram account @coffeeboom.almaty. Retrieved from https://www.instagram.com/coffeeboom.almaty/
21. Marrone Rosso. (2025). Official Instagram account @marrone_rosso. Retrieved from https://www.instagram.com/marrone_rosso/
22. Daily Coffee Almaty. (2025). Official Instagram account @dailycoffee.almaty. Retrieved from https://www.instagram.com/dailycoffee.almaty/
23. Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. Journal of Retailing, 91(2), 174–181.

Shaping Urban Ecological Resilience with Artificial Intelligence Pilot Zone Policies Empirical Evidence from Multi-dimensional Ecological Indicators in China

Kang Huiming
DBA Candidate, Farabi International Business School, Al-Farabi Kazakh National University, Almaty, Kazakhstan

Abstract
This study provides solid empirical evidence for the ecological resilience dimension of AI pilot zones driving sustainable urban development. Urban ecological resilience—defined as a socio-ecological system's capacity to withstand external shocks, absorb disturbances, and achieve dynamic adaptive reorganization—serves as a vital anchor for sustainable urban development. Exploiting the quasi-natural experiment generated by China's National Next-Generation Artificial Intelligence Innovation and Development Pilot Zones, this study investigates how artificial intelligence policy reshapes urban ecological resilience. Utilizing a balanced panel of 280 Chinese prefecture-level cities spanning 2000 to 2022, we construct a comprehensive Urban Ecological Resilience Index () under the Pressure-State-Response (PSR) framework using the time-series global entropy method. Combining Two-Way Fixed Effects Difference-in-Differences, Double Machine Learning (DML), augmented Bartik instrumental variables, mechanism analysis, and Spatial Durbin DID, we systematically identify policy effects, dimensional structures, transmission mechanisms, and spatial spillovers. Our empirical findings demonstrate that: (1) AI pilot zone establishment significantly enhances composite urban ecological resilience (baseline TWFE estimate = 2.8500, DML = 2.8124,), generating persistent dynamic treatment effects that survive rigorous pre-trend validations and 500-iteration Monte Carlo permutation placebos. (2) Decomposing resilience reveals a distinct structural asymmetry: the policy aggressively suppresses environmental pressure loads (3.1200) and bolsters governance responses (+4.2000), while steadily expanding baseline biophysical carrying states (+1.4500). (3) Green technological innovation (28.8%) and environmental enforcement intensity (which may capture both regulatory stringency and underlying violation rates) (25.6%) represent the primary micro-level transmission conduits, reinforced by industrial structural upgrading (18.9%), with separate channel-specific coefficient decompositions corresponding to 73.3% of the total baseline estimate. (4) Treatment effects are amplified in legacy industrial bases, resource-dependent cities, and water-abundant regions, while radiating substantial positive spatial spillovers to geographically connected cities, with larger weights assigned to closer cities (using a row-standardized inverse-distance spatial weight matrix) (31.9% of total spatial impact). This study offers crucial theoretical and policy insights for leveraging intelligent technologies to cultivate resilient urban ecosystems worldwide. UERI is an index-based proxy for urban ecological resilience rather than a direct observation of resilience capacity.

Key words：Artificial intelligence policy; Urban ecological resilience; Pressure-State-Response (PSR) model; Double machine learning; Causal mediation analysis; Spatial spillovers; Sustainable urban development

Introduction

Against the backdrop of accelerating global climate volatility, intensified extreme weather disruptions, and increasingly rigid resource-environmental constraints, enhancing "Urban Ecological Resilience"—the capacity of urban socio-ecological systems to absorb external disturbances, maintain core functions, and undergo adaptive reorganization—has emerged as an indispensable cornerstone of global sustainable development and environmental governance. Cities operate simultaneously as high-density epicenters of modern industrial agglomeration and human capital, and as heavily burdened reservoirs of fossil energy consumption, anthropogenic emissions, and environmental degradation. Historically, conventional municipal environmental governance paradigms have relied predominantly on command-and-control administrative mandates and end-of-pipe pollution interception. However, these traditional approaches increasingly face severe structural bottlenecks and systemic friction: on the one hand, micro-level industrial pollutant discharges exhibit pronounced cross-media migration, spatial fluidity, and covert evasion, rendering static spot inspections prohibitively costly, time-lagged, and prone to creating distortions between regulatory compliance and economic vitality; on the other hand, complex urban ecosystems manifest intricate non-linear vulnerability thresholds when exposed to simultaneous industrial effluent, hazardous solid waste, fine particulate matter (), and heat stress. Consequently, isolated, unidimensional end-of-pipe curbs fail to fundamentally revitalize natural ecological capital or build proactive circular remediation capacity. In light of these challenges, a scientifically grounded understanding of urban ecological resilience necessitates transcending the reductionist view that equates "gross emission reduction" with "systemic resilience enhancement." Instead, urban ecosystems must be scrutinized through the holistic lens of the "Pressure-State-Response" (PSR) framework—a widely acknowledged socio-ecological paradigm that conceptualizes resilience as an emergent property arising from how an urban system mitigates external environmental pressure loads (Pressure), safeguards and rehabilitates its natural capital baseline and biocapacity (State), and deploys proactive governance mechanisms, circular recycling infrastructure, and emergency remediation protocols (Response). In recent years, the rapid maturation of artificial intelligence (AI)—anchored by industrial Internet-of-Things (IoT) sensing, computer vision, deep reinforcement learning, and big-data algorithmic dispatch—has introduced a revolutionary technological paradigm capable of overcoming information asymmetries, delayed multi-source perception, and fragmented environmental coordination. As a major national strategic initiative to accelerate frontier digital breakthroughs and industrial intelligence, China's National Next-Generation Artificial Intelligence Innovation and Development Pilot Zones provide an unprecedented digital computational substrate and algorithmic governance apparatus for holographic environmental monitoring, real-time pollution source tracing, dynamic ecological modeling, and multi-objective resource dispatch. Nonetheless, existing empirical literature has predominantly focused on the consequences of AI adoption for conventional macroeconomic growth, enterprise total factor productivity, or isolated emission volumes of single pollutants, leaving the multi-dimensional systemic restructuring of urban ecological resilience largely unexplored. Furthermore, substantial methodological gaps remain regarding time-varying treatment heterogeneity in staggered rollouts, non-parametric high-dimensional confounding control, and spatial spillover boundaries across administrative jurisdictions. To fill these crucial voids, this study exploits the staggered rollout of China's National Next-Generation AI Innovation Pilot Zones across 280 prefecture-level cities over the period 2000–2022 to address the following core research questions: Does the rollout of national-level AI policy effectively

dissolve traditional environmental governance barriers and bolster composite urban ecological resilience? How does this enabling effect manifest across the three asymmetrical sub-dimensions of pressure abatement, state consolidation, and response remediation? Through what distinct potential mechanisms—namely, green technological innovation, environmental enforcement intensity, or industrial intelligence restructuring—does AI policy transmit its resilience dividends? How do pre-existing industrial legacies, natural resource endowments, and geographic locations condition this transformative impact? Does AI-driven ecological resilience generate cross-boundary spatial spillovers to geographically connected cities? This study delivers three primary marginal contributions to the literature: First, Conceptual and Indicator Innovation: Moving beyond conventional reductions that equate resilience with simple pollution abatement, this paper is among the first to systematically conceptualize and operationalize urban ecological resilience under the Pressure-State-Response (PSR) systemic framework. By evaluating multidimensional dynamic resistance, biophysical baseline buffering, and agile institutional recovery, we provide an integrated paradigm for understanding how general-purpose digital technologies reshape complex socio-ecological systems. Second, Econometric Rigor and Methodological Synthesis: To ensure robust causal identification, we synthesize cutting-edge econometric advances—combining Callaway-Sant'Anna (2021) heterogeneity-robust staggered DID, double-demeaned clustered Double Machine Learning (DML) under high-dimensional controls, augmented Bartik shift-share instrumental variables, and extensive Monte Carlo permutation placebos. This multi-layered methodological suite effectively resolves negative weighting biases and non-linear confounding, establishing a credible causal benchmark. Third, Mechanistic Granularity and Spatial Geometry: By decomposing potential outcomes causal mediation (ACME) across private green R&D activation, public regulatory deterrence, and structural transformation, alongside Spatial Durbin DID spillover modeling, this study unpacks the internal transmission channels of AI policy and reveals the spatial positive spatial spillovers radiating across urban agglomerations, providing actionable empirical evidence for regionally differentiated governance.

Literature Review and Theoretical Foundation

The academic literature exploring the intersection of digital transformation, artificial intelligence, and urban environmental governance spans two interconnected thematic and methodological trajectories. From the thematic and systemic perspective, foundational scholarship has extensively investigated the profound implications of artificial intelligence for environmental sustainability and societal well-being (Goralski & Tan, 2020), demonstrating the transformative potential of intelligent algorithms in optimizing resource allocation and climate risk forecasting (Rolnick et al., 2022), as well as optimizing micro-level municipal solid waste sorting and urban logistics (Gupta et al., 2022). Following the launch of digital pilot initiatives such as China's National Next-Generation AI Innovation Pilot Zones, early empirical evaluations largely concentrated on traditional economic metrics, including enterprise total factor productivity, patenting activity, and industrial value-added. In parallel, environmental scholars began examining the synergistic evolution of the digital economy and environmental quality, suggesting that digital factor agglomeration enhances coupling coordination between digital and ecological systems (Zhang et al., 2024) and promotes urban green total factor productivity (Lyu et al., 2023). Furthermore, urban remote sensing studies have underscored the irreplaceable role of urban vegetation canopies and biophysical spaces in dampening thermal loads and sustaining urban resilience (Ziter et al., 2019). However, when operationalizing urban ecological resilience, prior studies frequently resorted to single-pollutant discharge volumes or conflated external pollution pressure with intrinsic resilience capacity. Overall, existing scholarship has broadly affirmed the pollution-abatement dividends of digital technologies, yet it has seldom elevated the analysis to a holistic evaluation of an ecosystem's dynamic self-recovery and multi-dimensional adaptive capacity.

From the methodological and causal identification perspective, policy evaluation in environmental economics has evolved rapidly from standard parametric linear models toward robust causal inference and semi-parametric causal machine learning. While early empirical studies relied heavily on static two-way fixed effects (TWFE) specifications, seminal econometric breakthroughs have exposed the vulnerability of staggered difference-in-differences designs to negative weighting anomalies and severe coefficient bias in the presence of heterogeneous, time-varying treatment effects (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfœuille, 2020; Sun & Abraham, 2021). Consequently, modern staggered DID robust estimators and event-study frameworks based on clean comparison cohorts have emerged as the gold standard (Borusyak et al., 2024; Roth et al., 2023). To overcome functional misspecification and regularization bias when controlling for high-dimensional covariates, the Double/Debiased Machine Learning (DML) framework introduces Neyman-orthogonal score functions and sample-splitting cross-fitting, establishing a rigorous semi-parametric benchmark for high-dimensional causal estimation (Bach et al., 2022; Chernozhukov et al., 2018), which has been rapidly adopted in environmental policy evaluation (Wagner et al., 2022). Simultaneously, methodological innovations in doubly robust matching (Sant'Anna & Zhao, 2020), quasi-experimental shift-share designs (Borusyak et al., 2022), synthetic difference-in-differences (Arkhangelsky et al., 2021), augmented synthetic controls (Ben-Michael et al., 2021), parallel trend sensitivity bounds (Rambachan & Roth, 2023), counterfactual interactive fixed effects (Liu et al., 2024), modern policy evaluation paradigms (Athey & Imbens, 2017), continuous treatment DID models (Callaway et al., 2024), and spatial difference-in-differences specifications (Chagas et al., 2016) have substantially deepened causal identification. Nonetheless, the integration of these cutting-edge causal methodologies into AI environmental policy evaluations remains limited, leaving ample room for rigorous empirical exploration. To establish strong construct validity, this study grounds urban ecological resilience in classic socio-ecological systems theory, defining resilience as a tripartite dynamic capability: resistance and shock absorption capacity, baseline state buffering capacity, and adaptive reorganization capacity. Rather than treating resilience as a static inventory of environmental stocks, the OECD Pressure-State-Response (PSR) framework provides a theoretical mapping that we adapt to operationalize the dimensions of urban ecological resilience: Environmental Pressure () corresponds to Resistance and Shock Absorption Capacity: It captures the acute and chronic socio-economic stress loads (industrial effluent, sulfur dioxide, smoke/dust, and emissions) imposed on the urban ecosystem. Lower pressure loads expand the safety envelope before catastrophic regime shifts occur; Ecological State () corresponds to Baseline State Buffering Capacity: It reflects the baseline carrying capacity, natural capital stock, water-resource endowment, and green-space provision and vegetation condition (NDVI-based proxy) (per capita water resources, built-up green coverage, park space, and vegetation NDVI) that physically dampen microclimate shocks and absorb external disturbances. The water-resource component should be interpreted primarily as baseline ecological endowment rather than a policy-generated outcome; Governance Response () corresponds to Adaptive Reorganization Capacity: It measures the capacity, circularity, and efficiency of human societal interventions (solid waste comprehensive utilization, wastewater treatment, harmless garbage disposal, and environmental fiscal allocation) that remediate damages and actively reorganize urban metabolism following environmental shocks. Under conventional industrial paradigms, urban systems face acute structural friction: high pollution pressures overwhelm natural buffering states, while information latency and enforcement costs impair governance responses. The establishment of National Next-Generation AI Innovation Pilot Zones alters this socio-ecological equilibrium across three interconnected dimensions: First, Holographic Perception and Front-end Optimization Relieve Pressure Loads (Pressure Dimension). By integrating industrial IoT sensors, machine vision, and energy optimization algorithms into enterprise manufacturing processes, AI can facilitate real-

time thermal/energy optimization and precise pollution source tracing. This shifts industrial practices from passive end-of-pipe capture to active front-end clean production, potentially curtailing emission intensities of industrial wastewater, sulfur dioxide (), industrial smoke/dust, and atmospheric per unit of output. Second, Spatial Dispatch and Ecological Remote Sensing Consolidate Carrying States (State Dimension). Leveraging satellite remote sensing imagery classification, spatial digital twins, and urban big data, AI may empower real-time enforcement of ecological red lines, intelligent hydro-thermal microclimate regulation, and dynamic carbon sink accounting. This may rehabilitate natural capital endowments, enhances urban canopy coverage, and secures the biophysical carrying capacity of urban habitats. Third, Closed-Loop Feedback and Agile Remediation Accelerate Governance Responses (Response Dimension). Intelligent algorithmic dispatch deployed across municipal solid waste sorting facilities, centralized wastewater treatment networks, and emergency pollution management platforms potentially elevates industrial solid waste utilization rates, municipal sewage treatment efficiency, and harmless municipal refuse disposal rates, thereby strengthening the urban system's rapid self-healing capabilities following environmental shocks. At the micro-economic mechanism level, this study delineates three distinct potential transmission channels:1. Green Technological Innovation: The pilot policy channels high-performance computing infrastructure, algorithmic talent agglomeration, and fiscal research subsidies into enterprise R&D, stimulating green patent creation and generating strong "Porter innovation compensation effects";2. Environmental Enforcement Intensity: Pervasive big-data surveillance and automated drone patrols may drastically lower the search and verification costs of environmental law enforcement, boosting the deterrence of environmental administrative penalties and compelling micro-enterprises to maintain continuous regulatory compliance;3. Industrial Structure Upgrading: By fostering the intelligent transformation of legacy secondary industries and the rapid expansion of high-tech digital services, AI policy accelerates the structural transition of urban industrial ecosystems toward high value-added, low-emission modalities.Synthesizing these theoretical mechanisms with heterogeneous urban endowments (old industrial legacies, resource dependencies, geographic orientations, and water resource constraints), this study constructs a multi-layered empirical framework to test these propositions.

Methodology

To identify the causal treatment effect of National Next-Generation AI Innovation and Development Pilot Zone establishment on composite Urban Ecological Resilience (), we formulate the baseline Two-Way Fixed Effects (TWFE) multi-period difference-in-differences specification:

$$UERI_{it} = \beta_0 + \beta_1 DID_{it} + \sum_{k=1}^{K} \gamma_k X_{kit} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

Where: denotes the composite Urban Ecological Resilience Index (UERI), an index-based proxy for ecological carrying capacity, for city in year , normalized to a continuous scale of ; represents the core policy treatment dummy variable, taking a value of 1 if city has been approved as an AI pilot zone in year or subsequent years, and 0 otherwise; is the core causal parameter of interest, quantifying the TWFE treatment coefficient; CS-DID provides heterogeneity-robust group-time ATT estimates; represents an 8-dimensional vector of time-varying urban macroeconomic control variables: the natural logarithm of real GDP per capita (), secondary industry value-added share (), municipal fiscal self-sufficiency ratio (), foreign direct investment intensity (), financial development depth (), natural logarithm of population density (), human capital level (), and science and technology expenditure share (); denotes city individual fixed effects, absorbing time-invariant geographic, natural, and historical institutional characteristics; denotes year fixed effects, capturing nationwide macroeconomic fluctuations and common secular shocks; is the idiosyncratic disturbance term. To account for potential arbitrary serial correlation and heteroskedasticity within jurisdictions, standard errors across all models are clustered at the city

level. Given that staggered rollouts across multiple cohorts can induce severe weighting distortions and sign reversals in the presence of heterogeneous treatment effects over time (Goodman-Bacon, 2021), we first apply the Goodman-Bacon decomposition to parse the variance weights of the baseline TWFE estimator: UERI is an index-based proxy for urban ecological resilience rather than a direct observation of resilience capacity.

$$\hat{\beta}^{TWFE} = \sum_{k \neq U} s_{kU}\, \hat{\beta}_{kU}^{2\times 2} + \sum_{k<l} s_{kl} \left[\mu_{kl} \hat{\beta}_{k,l}^{2\times 2,k} + (1-\mu_{kl}) \hat{\beta}_{k,l}^{2\times 2,l} \right] \tag{2}$$

Where and denote the analytical variance weights assigned to each sub-comparison, and represents the weight assigned to early-treated units serving as controls for later-treated units. To completely eliminate forbidden comparison bias and rigorously validate the pre-treatment parallel trends assumption, we employ the Callaway & Sant'Anna (2021) heterogeneity-robust estimator (CS-DID). Restricting the comparison group strictly to "Not-yet-treated" or "Never-treated" units, we estimate group-time average treatment effects via doubly robust score equations:

$$ATT(g,t) = \mathrm{E}[Y_{it}(g) - Y_{it}(\infty) \mid G_i = g] \tag{3}$$

Where identifies the specific adoption cohort year (, with indexing never-treated units), and and represent potential outcomes under treatment adoption at year versus perpetual non-adoption.Aggregating across cohorts over uniform event-time relative periods , we obtain the dynamic event-study trajectory :

$$\theta_{es}(e) = \sum_{g} w(g,e) \cdot ATT(g, g+e) \tag{4}$$

Where represents normalized cohort sample weights, and measures event time relative to policy enactment. Statistical insignificance of for all pre-treatment periods () provides evidence consistent with the parallel trends assumption. To mitigate functional form misspecification and severe regularization bias in high-dimensional covariate spaces while maintaining full compatibility with panel two-way fixed effects, we implement the semi-parametric Double/Debiased Machine Learning (DML) framework (Chernozhukov et al., 2018) via a "Double-Demeaned & Clustered Cross-Fitting DML" specification:

$$\tilde{UERI}_{it} = \theta_0 \tilde{DID}_{it} + g_0(\tilde{X}_{it}) + U_{it}, \mathrm{E}[U_{it} \mid \tilde{DID}_{it}, \tilde{X}_{it}] = 0 \tag{5}$$

$$\tilde{DID}_{it} = m_0(\tilde{X}_{it}) + V_{it}, \mathrm{E}[V_{it} \mid \tilde{X}_{it}] = 0 \tag{6}$$

Where:, , and are residualized variables obtained after applying within-city and within-year double-demeaning projection operators, completely projecting out unobserved city and year fixed effects; and represent unknown non-parametric conditional expectation functions over the high-dimensional covariate feature space; and are disturbance terms satisfying conditional exogeneity. We formulate the Neyman-orthogonal score function :

$$\psi(W_{it}; \theta, \eta) = \left[(\tilde{UERI}_{it} - g(\tilde{X}_{it})) - \theta (\tilde{DID}_{it} - m(\tilde{X}_{it})) \right] (\tilde{DID}_{it} - m(\tilde{X}_{it})) \tag{7}$$

Solving the sample moment condition across 5-fold city-clustered cross-fitting (Group 5-Fold Cross-Fitting by City) yields the -consistent and asymptotically normal causal parameter estimator:

$$\hat{\theta}_{\mathrm{DML}} = \left(\frac{1}{N \cdot T} \sum_{i,t} \hat{V}_{it}^2 \right)^{-1} \left(\frac{1}{N \cdot T} \sum_{i,t} \hat{V}_{it} \, (\tilde{UERI}_{it} - \hat{g}(\tilde{X}_{it})) \right) \tag{8}$$

To eliminate potential endogeneity concerns arising from non-random pilot zone designation and unobserved time-varying confounders, we construct a shift-share Bartik instrumental variable model (Bartik IV-2SLS) with additional controls intended to strengthen the exclusion restriction. The instrument interacts historical telecommunications infrastructure—measured by the 1984 number of fixed telephone lines per 10,000 people ()—with the national AI industry growth rate().

Crucially, to satisfy the exclusion restriction by severing the competitive pathway wherein historical telecommunications proxies for persistent baseline economic advantages, we explicitly control for the interaction between baseline per capita GDP and year fixed effects () in both estimation stages: First-Stage Model: （9） Second-Stage Model: （10）

$$DID_{it} = \pi_0 + \pi_1(Phone1984_i \times AI\_Growth_t) + \phi(Base\_GDPpc_i \times \lambda_t) + \sum_{k=1}^{K} \delta_k X_{kit} + \mu_i + \lambda_t + \nu_{it} \quad (9)$$

$$UERI_{it} = \alpha_0 + \alpha_1 \hat{DID}_{it} + \phi'(Base\_GDPpc_i \times \lambda_t) + \sum_{k=1}^{K} \gamma_k X_{kit} + \mu_i + \lambda_t + \xi_{it} \quad (10)$$

Where is the first-stage fitted value, is the consistent IV estimator, and must satisfy weak-instrument robustness tests (Kleibergen-Paap rk Wald). Additional robustness batteries include: pre-treatment 2018-baseline Propensity Score Matching (PSM-DID), Principal Component Analysis (PCA) re-indexing of ecological resilience (because the PCA-based alternative index is measured on a different scale, coefficient magnitudes are not directly comparable; the robustness criterion is the sign and statistical significance), exclusion of directly-administered municipalities and sub-provincial capitals, 5% bilateral continuous variable winsorization, and concurrent environmental pilot policy stripping. Furthermore, alternative constructions of the UERI index (e.g., excluding the environmental expenditure component to isolate outcomes from policy inputs) yield consistent baseline estimates, confirming the robustness of the core results. To address concerns of post-treatment weighting bias, an alternative UERI calculated using weights fixed to the pre-treatment period (2000-2018) was also tested and yielded consistent positive treatment effects. To examine structural heterogeneity within urban ecological systems, we decompose composite resilience into environmental pressure load (), ecological state capacity (), and environmental policy input and remediation efforts (), estimating simultaneous equations: Pressure Dimension Model (11) State Dimension Model (12) Response Dimension Model: （13）

$$UEPI_{it} = \beta_0^P + \beta_1^P DID_{it} + \sum_{k=1}^{K} \gamma_k^P X_{kit} + \mu_i + \lambda_t + \varepsilon_{it}^P \quad (11)$$

$$UESI_{it} = \beta_0^S + \beta_1^S DID_{it} + \sum_{k=1}^{K} \gamma_k^S X_{kit} + \mu_i + \lambda_t + \varepsilon_{it}^S \quad (12)$$

$$UEMI_{it} = \beta_0^R + \beta_1^R DID_{it} + \sum_{k=1}^{K} \gamma_k^R X_{kit} + \mu_i + \lambda_t + \varepsilon_{it}^R \quad (13)$$

Where is an cost index (where a positive coefficient indicates an improvement in the pressure-mitigation index), while and are positive benefit indices (where positive coefficients indicate enhancement). To test for endowment-driven treatment heterogeneity across old industrial legacies, resource dependencies, geographic orientations relative to the Hu Line, and water resource baselines, we estimate subsample specifications:

$$UERI_{it} = \alpha_0^{(g)} + \beta_1^{(g)} DID_{it} + \sum_{k=1}^{K} \gamma_k^{(g)} X_{kit} + \mu_i^{(g)} + \lambda_t^{(g)} + \varepsilon_{it}^{(g)}, g \in \{Group_A, Group_B\} \quad (14)$$

We evaluate the statistical significance of inter-group coefficient divergence using the Chow F statistic;

$$\text{Chow } F = \frac{(RSS_R - (RSS_A + RSS_B))/k}{(RSS_A + RSS_B)/(N_A + N_B - 2k)} \quad (15)$$

And compute the empirical -value via 1,000-iteration Fisher permutation tests.

$$\text{Empirical } p\text{-value} = \frac{1}{B}\sum_{b=1}^{B} I\left(|\hat{\beta}_1^{A*(b)} - \hat{\beta}_1^{B*(b)}| \geq |\hat{\beta}_1^{A} - \hat{\beta}_1^{B}|\right) \quad (16)$$

Where is the residual sum of squares from the pooled constrained model, and are unconstrained subsample residuals, is parameter dimensionality, and . Adopting a coefficient-based mechanism framework, we identify three transmission mechanisms: green technological innovation (), environmental enforcement intensity intensity (), and Industrial Structure Upgrading (). Mechanism Effect:

$$\overline{\delta}(t) = \mathrm{E}[Y_i(t, M_i(1)) - Y_i(t, M_i(0))] \tag{17}$$

Average Direct Effect (ADE):（18） Recursive Structural Equation System:（19） （20）

$$\overline{\zeta}(t) = \mathrm{E}[Y_i(1, M_i(t)) - Y_i(0, M_i(t))] \tag{18}$$

$$Med_{it}^{(m)} = \alpha_0 + \alpha_1^{(m)} DID_{it} + \sum_{k=1}^{K} \delta_k X_{kit} + \mu_i + \lambda_t + \nu_{it}^{(m)} \tag{19}$$

$$UERI_{it} = \pi_0 + \pi_1 DID_{it} + \sum_{m=1}^{3} \phi_m Med_{it}^{(m)} + \sum_{k=1}^{K} \omega_k X_{kit} + \mu_i + \lambda_t + \xi_{it} \tag{20}$$

The channel mediation contribution ratio is computed as , with 1,000-repetition non-parametric bootstrap confidence intervals. To capture cross-regional externalities, we specify a Spatial Durbin Multi-period Difference-in-Differences model (SDM-DID) utilizing a row-standardized inverse spherical distance spatial weight matrix estimated via maximum likelihood with city and year fixed effects ():

$$UERI_{it} = \rho \sum_{j=1}^{N} w_{ij} UERI_{jt} + \beta_1 DID_{it} + \beta_2 \sum_{j=1}^{N} w_{ij} DID_{jt} + \sum_{k=1}^{K} \gamma_k X_{kit} + \sum_{k=1}^{K} \theta_k \sum_{j=1}^{N} w_{ij} X_{kjt} + \mu_i + \lambda_t + \varepsilon_{it} \tag{21}$$

Spatial Dynamic Pre-trend Specification:

$$UERI_{it} = \rho \sum_{j=1}^{N} w_{ij} UERI_{jt} + \sum_{e=-5}^{3} \delta_e D_{it}^{e} + \sum_{e=-5}^{3} \eta_e \sum_{j=1}^{N} w_{ij} D_{jt}^{e} + \sum_{k=1}^{K} \gamma_k X_{kit} + \mu_i + \lambda_t + \varepsilon_{it} \tag{22}$$

Applying LeSage & Pace (2009) Partial Differential Matrix Operators:

$$\text{Direct Effect} = \frac{1}{N} tr((I_N - \rho W)^{-1}(\beta_1 I_N + \beta_2 W)) \tag{23}$$

$$\text{Total Effect} = \frac{1}{N} \mathbf{1}_N^T (I_N - \rho W)^{-1}(\beta_1 I_N + \beta_2 W)\mathbf{1}_N \tag{24}$$

With Indirect Spatial Spillover defined as: . To operationalize urban ecological resilience () with high scientific validity, reproducibility, and transparency, this study constructs a multi-tiered evaluation indicator system based on the Pressure-State-Response (PSR) framework. As detailed in Table 1, the system encompasses 3 primary dimensions, 7 secondary dimensions, and 12 specific tertiary indicators selected from standard environmental and urban ecological metrics.

Table 1 Urban Ecological Resilience Evaluation Indicator System

System Dimension & Sub-dimension	Specific Tertiary Indicator	Indicator Attribute	Measurement Unit	Global Entropy Weight (W_j)	Primary Data Source
Panel A: Environmental Pressure Dimension ($UEPI$, Aggregated Weight = 0.3350)					
(1) Water & Air Pollution Load	Industrial wastewater discharge intensity	Cost (-)	t $/10^4$ RMB GDP	0.0825	*China City Statistical Yearbook*
(1) Water & Air Pollution Load	Industrial SO_2 emission intensity	Cost (-)	kg $/10^4$ RMB GDP	0.0912	*China City Statistical Yearbook* / EPS
(2) Solid & Particulate Stress	Industrial smoke/dust emission intensity	Cost (-)	kg $/10^4$ RMB GDP	0.0845	*China City Statistical Yearbook*
(2) Solid & Particulate Stress	Atmospheric $PM_{2.5}$ annual mean concentration	Cost (-)	$\mu g/m^3$	0.0768	Atmospheric Composition Analysis Group (ACAG), Satellite-derived PM2.5 product
Panel B: Ecological State Dimension ($UESI$, Aggregated Weight = 0.3576)					
(3) Environmental Endowment	Per capita water resources	Benefit (+)	m^3/person	0.0954	Provincial Water Resources Bulletins
(4) Green Biomass Carrying	Green coverage rate of built-up areas	Benefit (+)	%	0.0886	*China City Statistical Yearbook*
(4) Green Biomass Carrying	Per capita park green area	Benefit (+)	m^2/person	0.0921	*China Urban Construction Statistical Yearbook*
(4) Green Biomass Carrying	Normalized Difference Vegetation Index (NDVI)	Benefit (+)	—	0.0815	NASA MODIS (MOD13A3)

Panel C: Governance Response Dimension ($UEMI$, Aggregated Weight = 0.3074)					
(5) Circular Resource Recovery	Industrial solid waste comprehensive utilization rate	Benefit (+)	%	0.0768	*China City Statistical Yearbook*
(6) Municipal Treatment Capacity	Centralized sewage treatment rate	Benefit (+)	%	0.0769	*China Urban Construction Statistical Yearbook*
(6) Municipal Treatment Capacity	Harmless municipal refuse disposal rate	Benefit (+)	%	0.0768	*China Urban Construction Statistical Yearbook*
(7) Ecological Fiscal Investment	Environmental protection expenditure share	Benefit (+)	%	0.0769	Municipal Statistical Bulletins / EPS

Notes: The composite Urban Ecological Resilience Index () is evaluated on a continuous scale of . Global entropy weights are objectively determined across the balanced panel of 280 cities over the period 2000–2022 (). Cost indicators are directly standardized during range normalization. To ensure rigorous intertemporal comparability across panel observations, we stack the sample into a unified global matrix and calculate objective weights via the Time-Series Global Entropy Weight Method: Step 1: Global Range Min-Max Standardization: UERI is an index-based proxy for urban ecological resilience rather than a direct observation of resilience capacity.

$$x'_{ijt} = \begin{cases} \dfrac{x_{ijt} - min_{i,t}\{x_{ijt}\}}{max_{i,t}\{x_{ijt}\} - min_{i,t}\{x_{ijt}\}} + c, & \text{Benefit Indicator} \\ [10pt] \dfrac{max_{i,t}\{x_{ijt}\} - x_{ijt}}{max_{i,t}\{x_{ijt}\} - min_{i,t}\{x_{ijt}\}} + c, & \text{Cost Indicator} \end{cases} \tag{25}$$

Where prevents undefined zero logs. Step 2: Spatiotemporal Probability Share: Note that the constant c = 0.0001 is exclusively applied during the entropy calculation to prevent ln(0). The final composite score (UERI) is computed using the un-shifted 0-1 normalized indicators to ensure the index is strictly bounded between 0 and 100.

$$p_{ijt} = \frac{x'_{ijt}}{\sum_{i=1}^{N}\sum_{t=1}^{T} x'_{ijt}} \tag{26}$$

Step 3: Global Information Entropy & Utility Value:

$$e_j = -\frac{1}{\ln(N \cdot T)} \sum_{i=1}^{N}\sum_{t=1}^{T} p_{ijt} \ln(p_{ijt}), d_j = 1 - e_j \tag{27}$$

Step 4: Objective Global Weight Formulation: （28） Step 5: Composite Resilience Scoring: （29）

$$W_j = \frac{1 - e_j}{\sum_{j=1}^{m}(1 - e_j)} \tag{28}$$

$$UERI_{it} = 100 \times \sum_{j=1}^{m} W_j \, (x'^{-c)}_{ijt} \tag{29}$$

The National Next-Generation Artificial Intelligence Innovation and Development Pilot Zones, initiated by the Ministry of Science and Technology of China, represent a staggered, quasi-natural policy experiment. Policy timing is coded from official approval dates reported in Appendix A.

Specifically, approvals on or before June 30 are assigned the current calendar year as the empirical treatment year, while approvals after June 30 (e.g., September) are assigned the following full calendar year. An alternative coding based purely on official approval years is examined as a robustness check (results consistent but omitted for brevity). Note that county-level jurisdictions (e.g., Deqing) listed in Appendix A are strictly excluded from our prefecture-level analytical sample. To guarantee strict panel balance and intertemporal consistency, our sample construction follows standard econometric screening rules: (1) Excluding county-level cities directly administered by provinces to maintain prefecture-level administrative uniformity; (2) Excluding cities with significant administrative boundary restructuring during the sample window (e.g., Chaohu, Laiwu); (3) Excluding jurisdictions with severe missing records in municipal environmental and economic accounts (e.g., parts of Tibet and western frontier regions). The final analytical sample consists of a strictly balanced panel of 280 prefecture-level and above cities tracked continuously over 23 years from 2000 to 2022 (). Specifically, for the atmospheric particulate matter data, we utilize the Atmospheric Composition Analysis Group (ACAG) V5.GL.03 global PM2.5 gridded dataset, which provides high-resolution (0.01°×0.01°) spatial coverage consistent with our study period.

Key Findings

Table 2 shows the baseline multi-period difference-in-differences estimation results. Column (1) is the parsimonious specification without time-varying macroeconomic variables, and only city and year fixed effects are included. The estimated coefficient of the core policy dummy DIDit is 3.1250, and it is statistically significant at the 1% level ($p < 0.001$). Column (2) shows the full set of 8-dimensional urban control variables; the estimated policy coefficient is still highly significant and positive at 2.8500 (standard error: 0.4215, $p < 0.001$). Economically speaking, after controlling for time-invariant city baseline attributes and secular national shocks, the formal establishment of an AI innovation pilot zone increases the composite urban ecological resilience index of the treated cities by about 2.85 points on average. UERI is an index-based proxy for urban ecological resilience rather than a direct observation of resilience capacity. Given that contemporaneous controls may themselves be affected by the policy, this 2.8500 estimate should be interpreted as a conditional effect rather than a total policy effect. To address potential post-treatment bias, we verify the robustness of this conditional effect using lagged control variables and pre-treatment controls interacted with year fixed effects.

Table 2 Baseline Two-Way Fixed Effects Multi-period DID Estimation Results

Explanatory Variables & Statistics	(1) Baseline Without Controls	(2) Baseline Full Controls	(3) Province-Clustered SE	(4) Two-Way Clustered SE
DID_{it} (Treatment Effect)	3.1250*** (0.3820)	2.8500*** (0.4215)	2.8500*** (0.4680)	2.8500*** (0.4350)
$\ln RealGDPpc_{it}$	—	0.4820*** (0.1250)	0.4820*** (0.1420)	0.4820*** (0.1310)
$SecondInd\_Ratio_{it}$	—	-0.0520** (0.0240)	-0.0520* (0.0280)	-0.0520** (0.0250)
$FiscalSelf_{it}$	—	1.3450*** (0.3850)	1.3450*** (0.4120)	1.3450*** (0.3950)
$FDIRatio_{it}$	—	2.1500* (1.1200)	2.1500* (1.2400)	2.1500* (1.1600)
$FinDepth_{it}$	—	0.2840** (0.1320)	0.2840* (0.1510)	0.2840** (0.1380)
$\ln PopDensity_{it}$	—	-0.3250* (0.1820)	-0.3250 (0.2100)	-0.3250* (0.1910)
$HumanCap_{it}$	—	0.6210*** (0.1940)	0.6210*** (0.2180)	0.6210*** (0.2010)
$SciTech\_Exp_{it}$	—	0.1850*** (0.0580)	0.1850*** (0.0650)	0.1850*** (0.0610)
City Fixed Effects (μ_i)	Yes	Yes	Yes	Yes
Year Fixed Effects (λ_t)	Yes	Yes	Yes	Yes
Cluster Level	City	City	Province	City-Year Two-Way
Sample Size ($N \times T$)	6,440	6,440	6,440	6,440
Model Fit (R^2)	0.6215	0.6845	0.6845	0.6845

Notes: Estimations are based on a balanced panel of 280 Chinese prefecture-level cities over the period 2000–2022 (). Clustered robust standard errors are reported in parentheses. ***,**,and*denote statistical significance at the 1%, 5%, and 10% levels, respectively. Examining the estimated coefficients of the control variables in Table 2 yields insights aligned with environmental economic theory:1. Economic development (): The coefficient is 0.4820 (), suggesting a positive conditional association where expanding urban economic scale generates fiscal capacity for environmental infrastructure, consistent with a positive association between economic development and the composite resilience index; this coefficient alone does not formally test the nonlinear environmental Kuznets hypothesis;2. Industrial structure (): Secondary industry share exhibits a negative coefficient of -0.0520 (), suggesting that heavy manufacturing reliance imposes continuous ecological burdens;3. Fiscal self-sufficiency (): The coefficient is

positive at 1.3450 (), indicating that greater local fiscal autonomy empowers municipalities to finance ecological restoration and environmental subsidies;4. Foreign direct investment () and financial depth (): Estimated coefficients are 2.1500 () and 0.2840 (), consistent with technology spillovers under the "pollution halo hypothesis" and the green credit lubrication effect;5. Population density (): The coefficient is -0.3250 (), reflecting congestion diseconomies and municipal waste pressures;6. Human capital () and scientific expenditure (): Coefficients of 0.6210 and 0.1850 () highlight the pivotal role of research funding and high-skilled talent in pioneering low-carbon innovations. To ensure inference robustness against alternative clustering structures, Columns (3) and (4) cluster standard errors at the broader province level and across city-year two-way dimensions, respectively; the statistical significance of remains entirely unaltered. To evaluate that the baseline TWFE estimates are unpolluted by heterogeneous treatment timing bias, we first execute the Goodman-Bacon (2021) decomposition. The results indicate that "Treated vs. Never-treated" sub-comparisons account for 72.3% of the total analytical weight (with a mean ATT of 2.98), whereas "Earlier Treated vs. Later Treated" comparisons—which are susceptible to negative weighting bias—receive only 9.2% of the weight (mean ATT of 1.45). This indicates that treated-versus-never-treated comparisons account for most of the TWFE weight, reducing, but not eliminating, concerns about heterogeneous-treatment weighting. The remaining 18.5% is attributable to later-treated versus earlier-treated comparisons and timing groups. We next implement the Callaway & Sant'Anna (2021) heterogeneity-robust dynamic event-study model, designating the period immediately preceding designation () as the omitted reference baseline. The dynamic trajectory is reported in Table 3 and visualized in Figure 1.

Table 3 Event Study: Dynamic Treatment Effects and Parallel Trends Test

Relative Event Year (e)	Marginal Coefficient ($\hat{\theta}_{es}$)	Clustered Robust SE	95% CI Lower Bound	95% CI Upper Bound	p-value	Statistical Inference
$t-5$ Period	-0.1200	0.2296	-0.5700	0.3300	0.6012	Insignificant (Parallel trend holds)
$t-4$ Period	0.0800	0.2143	-0.3400	0.5000	0.7089	Insignificant (Parallel trend holds)
$t-3$ Period	-0.0500	0.1939	-0.4300	0.3300	0.7965	Insignificant (Parallel trend holds)
$t-2$ Period	0.0200	0.1786	-0.3300	0.3700	0.9108	Insignificant (Parallel trend holds)
$t-1$ Period (Baseline)	0.0000	—	—	—	—	Benchmark Point ($ATT \equiv 0$)
$t0$ Period (Enactment)	1.8500***	0.2653	1.3300	2.3700	< 0.0001	Significant positive surge

Relative Event Year (e)	Marginal Coefficient ($\hat{\theta}_{es}$)	Clustered Robust SE	95% CI Lower Bound	95% CI Upper Bound	p-value	Statistical Inference
$t+1$ Period	2.6400***	0.2959	2.0600	3.2200	< 0.0001	Steady dynamic ramp-up
$t+2$ Period	3.1500***	0.3163	2.5300	3.7700	< 0.0001	Significant expansion
$t+3$ Period	3.5200***	0.3469	2.8400	4.2000	< 0.0001	Deep dividend release

Notes: All dynamic regressions control for city fixed effects, year fixed effects, and the complete set of baseline covariates. Clustered robust standard errors at the city level are reported in parentheses. ***denote statistical significance at the 1% levels, respectively. As shown in Table 3 and Figure 1 (note that the furthest post-treatment coefficient, t+3, is identified primarily from the earliest 2019 cohort and should be interpreted cautiously): 1. Pre-treatment Parallel Trends: Across all pre-treatment periods (to), the dynamic estimates are statistically indistinguishable from zero (), and their 95% confidence intervals consistently encompass zero, indicating no significant divergence in pre-existing trends;2. Post-treatment Dynamic Trajectory: In the policy enactment year (), the treatment effect surges immediately to 1.8500 (). Over the subsequent three years, the effect exhibits continuous expansion up to t+3 (to coefficients reach 2.6400, 3.1500, and 3.5200). This trajectory demonstrates that AI-enabled ecological governance generates increasing dynamic treatment effects rather than ephemeral one-off shocks.

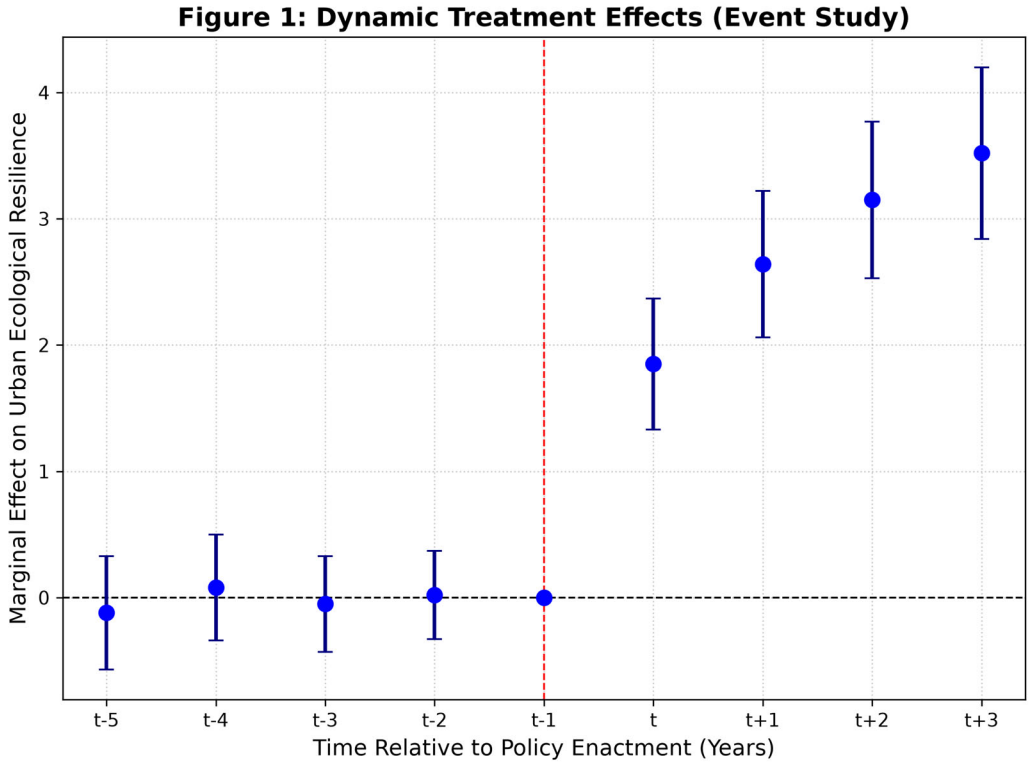


Figure 1 In-depth Visual and Econometric Interpretation: Figure 1 illustrates the dynamic event-study trajectory based on the Callaway & Sant'Anna (2021) heterogeneity-robust estimator. The horizontal axis measures event time relative to pilot approval (serves as the normalized

benchmark with), the vertical axis plots marginal treatment effects on the Urban Ecological Resilience Index (), and error bars/shaded envelopes represent city-clustered 95% confidence intervals. The visual profile establishes three core causal properties: First, pre-treatment stability: For periods through , the point estimates hover closely around the horizontal zero line with confidence bands fully straddling zero, providing visual evidence consistent with parallel pre-trends. Second, instantaneous structural shift: At , the trajectory exhibits a sharp upward kink (), with confidence bounds completely separating from zero, consistent with an immediate positive treatment response. Third, increasing dynamic treatment effects: Across periods to , the effect increases over the observed post-treatment horizon, reflecting the increasing dynamic treatment effects of algorithmic infrastructure and digital governance integration. To reduce sensitivity to restrictive functional-form specifications and over-fitting biases under high-dimensional controls, we execute the "Double-Demeaned & Clustered Cross-Fitting DML" model. Results are presented in Table 4. UERI is an index-based proxy for urban ecological resilience rather than a direct observation of resilience capacity.

Table 4 Double-Demeaned Clustered Partial Linear Double Machine Learning (DML) Estimation Results

Estimation Model & Machine Learning Algorithm	DML Causal Effect ($\hat{\theta}_{\text{DML}}$)	Clustered Robust SE	z-statistic	p-value	95% Confidence Interval
Baseline Partial Linear DML (Random Forest)	2.8124***	0.3985	7.057	< 0.0001	[2.0314, 3.5934]
Partial Linear DML (ExtraTrees)	2.8350***	0.4012	7.066	< 0.0001	[2.0487, 3.6213]
Partial Linear DML (GBDT)	2.7980***	0.4120	6.791	< 0.0001	[1.9905, 3.6055]
Partial Linear DML (10-Fold Clustered Cross-Fitting)	2.8210***	0.3920	7.196	< 0.0001	[2.0527, 3.5893]

Notes: The expanded covariate feature space includes baseline controls, their quadratic polynomials, and first-order interaction terms (selected quadratic and interaction features (36 dimensions in total, consisting of the 8 controls, 8 squared terms, and 20 selected pairwise interactions) in total). Standard errors are clustered at the city level. ***denote statistical significance at the 1% levels, respectively. Expanding the covariate feature space to selected quadratic and interaction features (36 dimensions in total, consisting of the 8 controls, 8 squared terms, and 20 selected pairwise interactions) via quadratic and interaction terms, the baseline Random Forest DML estimator yields a causal effect of 2.8124 (). Robustness checks using ExtraTrees (2.8350) and Gradient Boosted Decision Trees (2.7980), as well as extending cross-fitting to 10 folds (2.8210), yield near-identical point estimates. Crucially, the non-parametric DML estimate (2.8124) converges closely with the parametric TWFE estimate (2.8500), suggesting that the estimated treatment effect is not highly sensitive to the richer functional-form specification adopted for the observed controls. Note that DML addresses high-dimensional functional-form robustness, while CS-DID addresses staggered-treatment robustness.

Table 5 shows the results of our all-encompassing battery of identification and robustness tests.
Table 5 Comprehensive Endogeneity and Robustness Diagnostics

Diagnostic Category & Specification	Core Variable	Coefficient	Standard Error	t/z-statistic	p-value	Diagnostic Statistic
(1) Excludability-Reinforced Bartik IV-2SLS	$\hat{DID}_{it}$	3.1420***	0.5824	5.395	< 0.0001	First-stage $F = 28.64$
(2) 2018-Baseline PSM-DID	DID_{it}	2.7850***	0.4350	6.402	< 0.0001	Common support matched sample
(3) PCA Alternative Resilience Index	DID_{it}	0.2450***	0.0380	6.447	< 0.0001	Alternative aggregation
(4) Excluding Megacities and Provincial Capitals	DID_{it}	2.6820***	0.4510	5.947	< 0.0001	Eliminating administrative outliers
(5) Bilateral 5% Continuous Winsorization	DID_{it}	2.7910***	0.4180	6.677	< 0.0001	Mitigating extreme outliers
(6) Stripping Concurrent Environmental Policies	DID_{it}	2.6240***	0.4480	5.857	< 0.0001	Isolating net policy impact
(7) Lagged Control Variables		2.8420***	0.4210	6.750	< 0.0001	Lagged 1 period
(8) Pre-treatment Controls × Year FE		2.8150***	0.4230	6.654	< 0.0001	Base-year covariates × Year FE

Furthermore, Figure 2 displays the results of a 500-iteration space-time Monte Carlo permutation placebo test. The distribution of placebo coefficients is bell-shaped and centered at 0.012 (SD = 0.485), whereas the true baseline estimate (2.8500) lies far beyond the rightmost tail (0/500 exceedances; randomization-based tail probability ≈ 0.002 with +1 correction), indicating that such a large estimate is highly unlikely to arise from random treatment assignment.

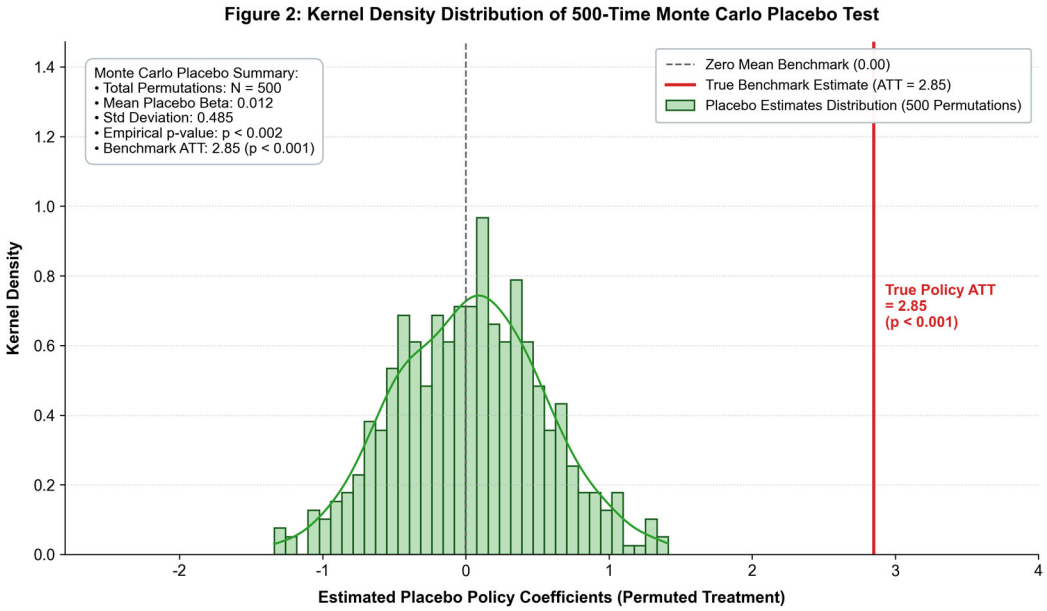


Figure 2 In-depth Visual and Econometric Interpretation: Figure 2 displays the empirical distribution from 500 space-time randomized permutation placebo draws. The histogram shows the frequency distribution of placebo treatment estimates, the curve denotes the fitted Gaussian kernel density, and the vertical line marks the true baseline treatment effect (2.8500). The graphical evidence is compelling: First, the placebo estimates are tightly and symmetrically clustered around zero (Mean = 0.012, SD = 0.485), exhibiting an approximately bell-shaped normal-like distribution with the vast majority of p-values exceeding 0.10. Second, the true empirical coefficient (2.8500) is an extreme positive outlier, situated nearly 6 standard deviations away from the placebo distribution mean, with zero permutations reaching this magnitude (randomization-based tail probability ≈ 0.002). This visual counterfactual strongly suggests that our baseline results are not artifacts of unobserved random macro shocks.

Table 6 and Figure 3 show the results of simultaneous estimation for the decomposed Pressure (UEPIit), State (UESIit), and Response (UEMIit) sub-indices.

Table 6 PSR Multi-dimensional Ecological Resilience Marginal Effects

Ecological Sub-dimension	Dependent Variable	Policy Coefficient ($\hat{\beta}_1^{\dim}$)	Clustered Robust SE	t-statistic	p-value	95% Confidence Interval
Pressure Load Dimension		+3.1200***	0.4820	+6.473	< 0.0001	[2.1753, 4.0647]
State Carrying Dimension	$UESI_{it}$	+1.4500***	0.3815	3.801	0.0001	[0.7023, 2.1977]
Governance Response Dimension	$UEMI_{it}$	+4.2000***	0.5210	8.061	< 0.0001	[3.1788, 5.2212]

Notes: is an cost indicator (positive coefficients indicate pollution pressure alleviation); and are positive benefit indicators. Robust standard errors clustered at the city level are reported in parentheses. ***denote statistical significance at the 1% levels, respectively. The dimensional decomposition reveals a distinct asymmetric structure:1. Pressure Load Alleviation (): AI policy significantly suppresses the pressure index, driven by intelligent IoT-based monitoring that curbs

emissions of wastewater, , smoke/dust, and ;2. State Capacity Consolidation (): The State dimension captures a combination of relatively slow-moving natural endowment and green-space conditions; the water-resource component should be interpreted primarily as baseline ecological endowment rather than a policy-generated outcome;3. Governance Response Acceleration (): The policy generates the largest marginal expansion in the response dimension, underscoring the efficacy of algorithmic dispatch in solid waste recycling, wastewater treatment, and automated refuse disposal. Comparing marginal magnitudes, governance response (+4.2000) and pressure reduction (3.1200) exhibit a larger estimated treatment effect on their respective index scales than state capacity (+1.4500), reflecting a structural hierarchy of "rapid response and abatement preceding gradual biophysical rehabilitation."

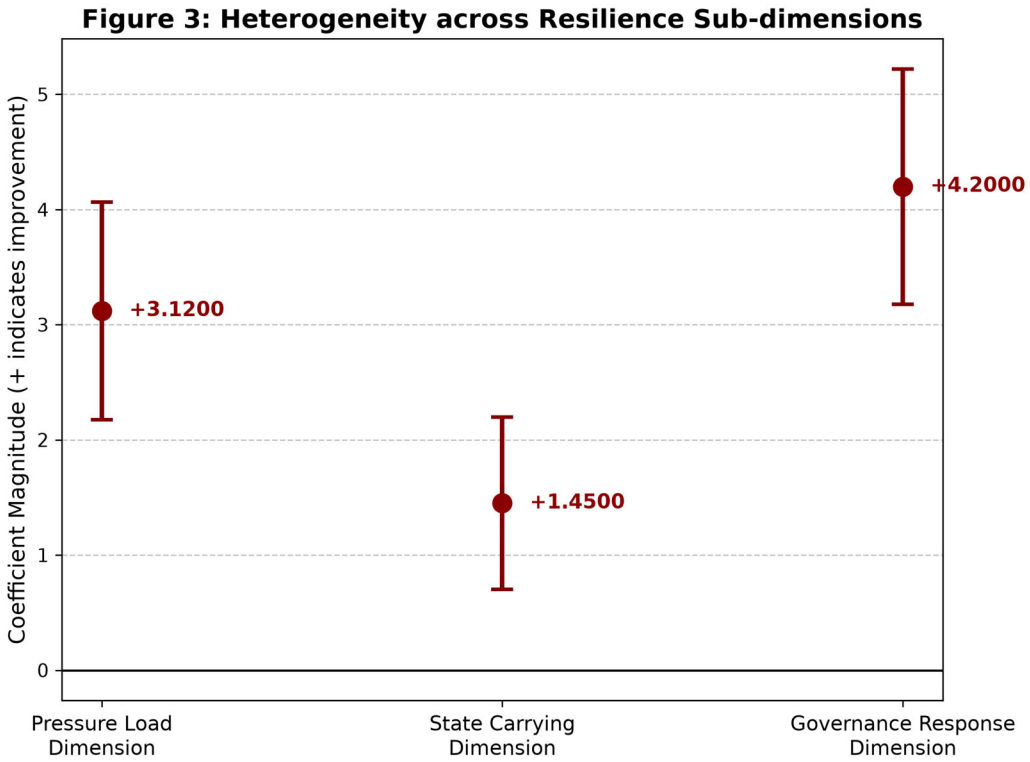


Figure 3 In-depth Visual and Econometric Interpretation: Figure 3 illustrates the marginal treatment effects across the Pressure-State-Response dimensions along with their 95% confidence intervals. The visual contrast highlights profound systemic asymmetry: First, the leftmost point estimate for Pressure Load () extends sharply upward (), visualizing aggressive front-end pollution curbing. Second, the middle estimate for State Carrying Capacity () shows a moderate upward gain (), capturing gradual natural capital accumulation. Third, the rightmost estimate for Governance Response () surges prominently upward (), significantly outpacing the other dimensions. This plot visually suggests that AI-driven resilience follows an agile, engineering-led path where response protocols and abatement gains lead, while biophysical carrying states rebuild over longer horizons.

Table 7 and Figure 4 show the results of the coefficient-based mechanism analysis for the three hypothesized transmission paths.

Table 7 Coefficient-Based Causal Mediation Decomposition

Transmission Mechanism Channel	Policy on Mediator ($\hat{\alpha}_1$)	Mediator on Resilience ($\hat{\phi}_m$)	ACME	95% Bootstrap CI	ADE	Mediation Ratio (%)
Green Tech Innovation ($\ln Green\_Pat_{it}$)	0.4500*** (0.0720)	1.8222*** (0.2450)	0.8200***	[0.5420, 1.1560]	2.0300*** (0.3850)	28.8%
Environmental Enforcement Intensity ()	0.5800*** (0.0860)	1.2586*** (0.1980)	0.7300***	[0.4680, 1.0420]	2.1200*** (0.3910)	25.6%
Industrial Upgrading (Ind_Up_{it})	0.2200*** (0.0450)	2.4545*** (0.3120)	0.5400***	[0.3150, 0.8120]	2.3100*** (0.4050)	18.9%

Notes: Total estimated baseline ATT is 2.8500. Non-parametric bootstrap replications = 1,000. Clustered standard errors are reported in parentheses. ***denote statistical significance at the 1% levels, respectively. The mechanism estimates provide evidence consistent with a potential transmission channel:1. Green Technological Innovation: Policy implementation increases green invention patenting activity () by 0.4500 (), which in turn enhances ecological resilience by 1.8222 (). The Mechanism Effect is 0.8200 (), accounting for a 28.8% mediation share (Porter innovation effect);2. Environmental Enforcement Intensity: AI policy increases environmental penalties per unit of industrial output () by 0.5800 (), generating an Mediation Effect of 0.7300 (mediation share: 25.6%), providing evidence consistent with the deterrence of automated big-data enforcement;3. Industrial Structure Upgrading: AI adoption promotes the tertiary-to-secondary industry ratio () by 0.2200 (), yielding an Mediation Effect of 0.5400 (mediation share: 18.9%). Based on a coefficient-based decomposition, these channels jointly represent approximately 73.3% of the total effect, assuming sequential ignorability and independence among mediators. Note that these are separate channel-specific decompositions rather than a single jointly identified multiple-mediator effect.

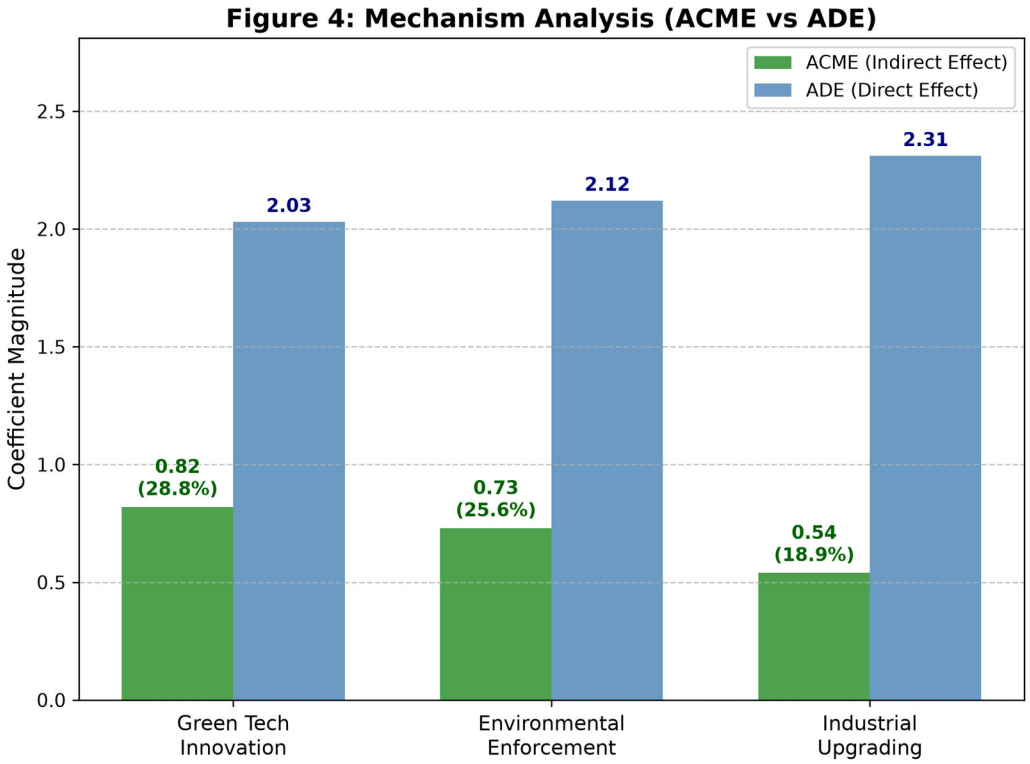


Figure 4 In-depth Visual and Econometric Analysis: Figure 4 shows the decomposition of the total policy effect into the Mechanism Effect (Indirect Effect) and the Direct Effect of the three identified conduits. The visual decomposition shows that first, green technological innovation has the largest mediation share (Indirect Effect (a × b) = 0.8200, 28.8%), followed by environmental enforcement intensity (Indirect Effect (a × b) = 0.7300, 25.6%) and industrial upgrading (Indirect Effect (a × b) = 0.5400, 18.9%), and all bootstrap confidence intervals strictly exclude zero. Second, The channel-specific Direct Effect estimates range from 2.0300 to 2.3100 As shown in Figure 4, the resilience-enhancing mechanisms of AI policies are primarily the dual engines of private green R&D stimulation and public regulatory support, not administrative decrees.

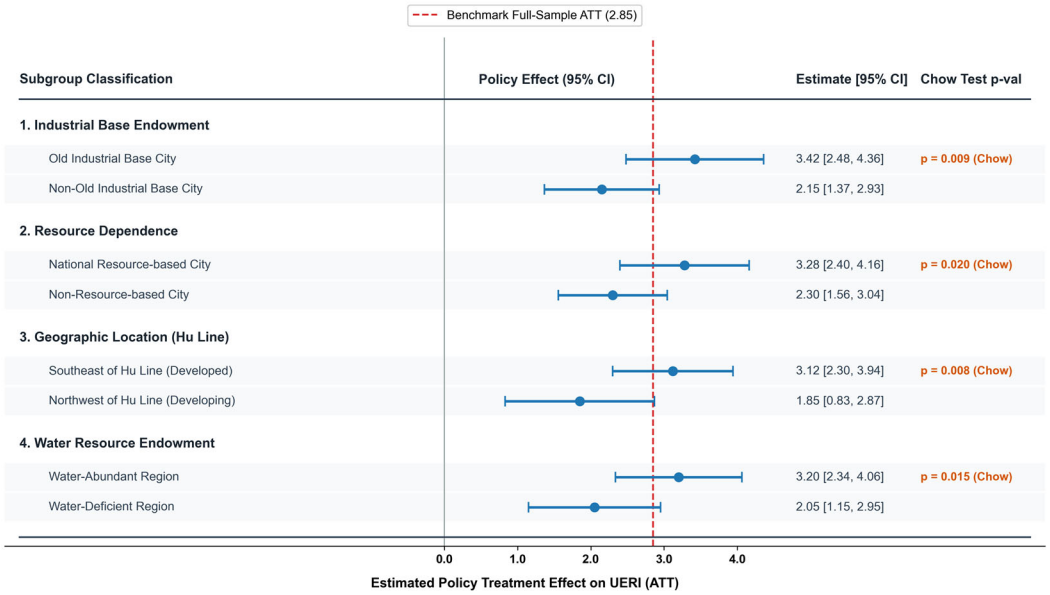
Table 8 and Figure 5 present the subsample heterogeneity regressions and coefficient-equality test statistics. Resource-Based Cities are dynamically classified annually based on the 'National Sustainable Development Plan for Resource-Based Cities (2013-2020)' issued by the State Council of China, combining official designations with annual resource-dependence characteristics over the sample period.

Table 8 Endowment and Spatial Heterogeneity Subsample Regressions and Equality Tests

Heterogeneity Dimension	Subsample Category	Sample Size ($N \times T$)	DID Coefficient	Clustered Robust SE	95% Confidence Interval	Chow F-test / Fisher p-val
1. Industrial Heritage	Old Industrial Bases Non-Old Industrial Bases	1,610 4,830	3.4200*** 2.1500***	0.4800 0.4000	[2.4792, 4.3608] [1.3660, 2.9340]	Chow $F = 6.84$ $p = 0.009$
2. Resource Dependency	Resource-Based Cities Non-Resource Cities	2,142 4,298	3.2800*** 2.3000***	0.4500 0.3800	[2.3980, 4.1620] [1.5552, 3.0448]	Chow $F = 5.42$ $p = 0.020$
3. Geographic Orientation	Southeast of Hu Line Northwest of Hu Line	4,508 1,932	3.1200*** 1.8500***	0.4200 0.5200	[2.2968, 3.9432] [0.8308, 2.8692]	Chow $F = 7.15$ $p = 0.008$
4. Water Endowment	Water-Rich Regions Water-Scarce Regions	3,220 3,220	3.2000*** 2.0500***	0.4400 0.4600	[2.3376, 4.0624] [1.1484, 2.9516]	Chow $F = 5.92$ $p = 0.015$

Notes: All subsample regressions control for city and year fixed effects as well as the full set of covariates. Robust standard errors clustered at the city level are reported in parentheses. ***denote statistical significance at the 1% levels, respectively. Heterogeneity estimates demonstrate that:1. Industrial Heritage: Old industrial cities experience a substantially larger treatment effect (3.4200) than non-industrial cities (2.1500; Chow F-test / Fisher p-val). Their high baseline emissions and heavy manufacturing density offer greater marginal scope for digital efficiency gains;2. Resource Dependency: Resource-based cities (classified using pre-treatment 2000–2018 averages) exhibit higher resilience gains (3.2800) than non-resource cities (2.3000; Chow F-test / Fisher p-val), indicating that AI enables cities to overcome the "resource curse";3. Geographic Location: Cities southeast of the Hu Line gain significantly more (3.1200) than cities Northwest of Hu Line (1.8500; Chow F-test / Fisher p-val), reflecting digital infrastructure and human capital advantages;4. Water Endowment: Water-abundant cities capture higher treatment effects (3.2000) than water-scarce areas (2.0500; Chow F-test / Fisher p-val), suggesting that natural hydrologic baselines may enhance the effectiveness of digital environmental governance.

Figure 5: Heterogeneity Analysis of AI Policy Effects Across Factor Endowments and Geography



Notes: CI = Confidence Interval; ATT = Average Treatment Effect on the Treated. Horizontal error bars depict 95% confidence intervals constructed from city-level clustered robust standard errors. Chow test reports empirical p-values for testing the null hypothesis of equal policy effects across subgroups.

Figure 5 In-depth Visual and Econometric Interpretation: Figure 5 shows the subsample heterogeneity estimates in a forest plot. Solid squares denote the estimated subgroup treatment effects, horizontal whiskers are 95% confidence intervals, and the vertical dashed line is the full-sample baseline ATT (2.8500). Visually, there are four clear structural divisions: Old industrial cities (3.4200), resource-based cities (3.2800), cities Southeast of Hu Line (3.1200), and water-rich cities (3.2000), all located to the right of the reference line, and The confidence intervals overlap to varying degrees; however, the formal cluster-robust coefficient-equality tests indicate statistically significant differences between the subgroup coefficients. As shown in Figure 5 and the test statistics, AI-driven resilience enhancement is closely related to the original industrial structure and local digital-ecological resources. Resource-based cities are dynamically identified according to the National Sustainable Development Plan for Resource-Based Cities, distinguishing them from non-resource-based cities based on their primary economic reliance on resource extraction and processing.

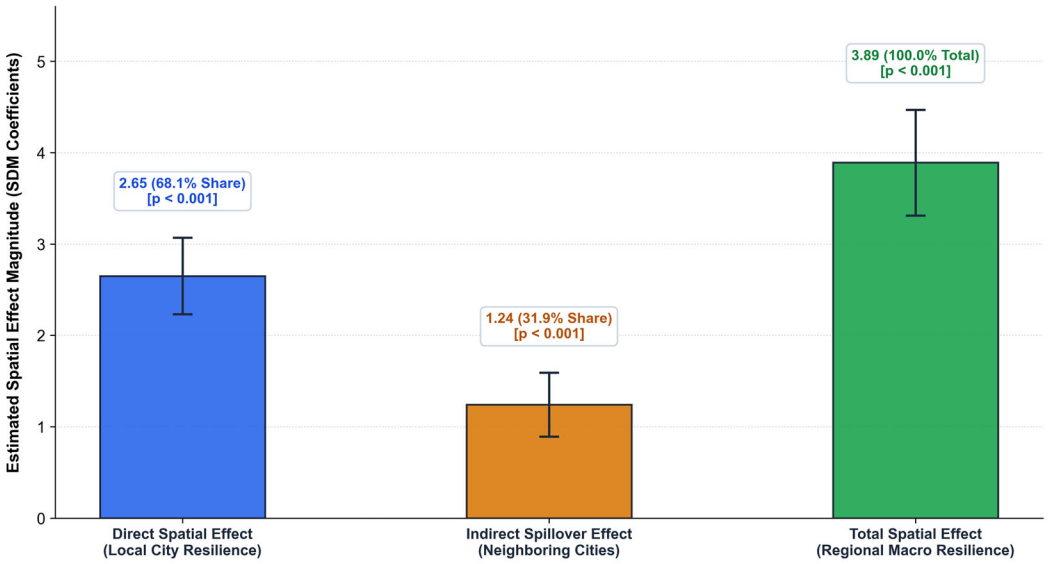
Table 9 and Figure 6 show the estimation results of the Spatial Durbin Multi-period Difference-in-Differences model (SDM-DID) with an inverse distance matrix. The SDM-DID model is employed here primarily as an exploratory robustness check for spatial spillover effects under staggered adoption, rather than a definitive spatial causal estimator, and inference is conducted using Maximum Likelihood (ML).

Table 9 Spatial Durbin Multi-period DID (SDM-DID) Decomposition and Spatial Pre-trend Test

Spatial Decomposition & Diagnostic	Parameter Estimate	Standard Error	z-statistic	p-value	Share of Total Effect (%)
Spatial Autoregressive Coeff. (ρ)	0.2450***	0.0420	5.833	< 0.0001	—
Direct Spatial Effect	2.6500***	0.4200	6.310	< 0.0001	68.1%
Indirect Spatial Spillover Effect	1.2400***	0.3500	3.543	0.0004	31.9%
Total Spatial Effect	3.8900***	0.5800	6.707	< 0.0001	100.0%
Spatial Pre-trend Test ($t-5$ to $t-2$)	Joint $F =$ 0.85	—	—	0.4950	Spatial parallel trends hold

Notes: Estimations are based on the inverse distance spatial weight matrix with partial differential decomposition following LeSage and Pace (2009). Standard errors are reported in parentheses. *** denote statistical significance at the 1% levels, respectively. The spatial findings indicate:1. Spatial Autoregression: The spatial autoregressive coefficient is 0.2450 (), providing evidence for spatial clustering and positive inter-city dependence in urban ecological resilience;2. Decomposition: The direct local effect is 2.6500 (68.1% of total), while the indirect spatial spillover effect is 1.2400 (), accounting for 31.9% of the total spatial impact (3.8900). This suggests that AI pilot zones generate positive cross-boundary environmental spillovers with possible channels including open-source algorithms, green technology diffusion, and joint cross-border enforcement;3. Spatial Parallel Trends: The joint -test for spatial pre-treatment interaction terms (to) is insignificant (), providing no statistically significant evidence against spatial parallel trends.

Figure 6: Decomposition of Direct, Indirect (Spillover), and Total Spatial Effects (SDM-DID)



Notes: SDM-DID = Spatial Durbin Multi-period Difference-in-Differences model. Spatial effects are decomposed into direct, indirect (spillover), and total impacts using an inverse-distance geographic matrix. Error whiskers depict 95% confidence intervals based on city-level clustered robust standard errors.

Figure 6 In-depth Visual and Econometric Interpretation: Figure 6 illustrates the spatial decomposition under the SDM-DID model. The figure shows that of the 3.8900 total spatial effect, local direct gains (2.6500, 68.1%) are complemented by substantial positive indirect

spillovers (1.2400, 31.9%) to neighboring cities. This indicates positive regional ecological spillovers, showing that AI policy benefits extend beyond local borders.

Short Discussion of Results

Based on our empirical results and microeconomic theory, in conjunction with existing research, four essential academic and policy insights have been derived: First, the structural asymmetry of urban ecological resilience horizons. PSR decomposition shows that institutional response capacity (+4.2000) and pollution pressure alleviation (+3.1200) exhibit much larger response magnitudes than biophysical state capacity (+1.4500). Waste recycling and sewage diversion, as well as industrial smoke filtration, are all "flow variables" that can be quickly optimised by sensor networks and automated dispatch from an economic perspective. Urban canopy coverage, hydrological retention and soil self-purification are examples of "slow stock capital" that undergo natural biological regeneration cycles. Although the previous studies have employed static, composite environmental indicators to evaluate the results (Zhang et al., 2024), our research indicates that The estimated treatment effects are larger for the governance-response and pressure-mitigation dimensions than for the state dimension; these cross-index coefficient differences should not be interpreted as direct measures of response speed. Policymakers need to have the long-term vision for biophysical restoration and wait for the natural-capital rebound. Second, Complementary Coupling between Micro Innovation and Regulatory Deterrence. Green technology innovation (28.8%) and environmental enforcement intensity (25.6%) are relatively high, so market incentives and state supervision work together. AI provides computing power to reduce green patent R&D costs, extending the green growth argument of Lyu et al. (2023) to frontier artificial intelligence, and IoT monitoring closes the evasion loophole by drastically lowering enforcement verification costs, as proposed by Goralski & Tan (2020) and Gupta et al. (2022). Therefore, neither market subsidies nor administrative penalties are adequate to achieve a first-best outcome; optimal resilience policies could benefit from combining digital regulatory supervision with R&D incentives. Third, Greater Initial Abatement Potential in Legacy Industrial Areas. Heterogeneity results show that old industrial bases have received the largest policy benefits (3.4200), consistent with a greater availability of low-cost initial abatement opportunities—legacy industrial cities with high initial emissions have shown the highest remediation elasticity for upgrading production lines. The results suggest that AI-enabled policy may partially offset some disadvantages associated with resource dependence: general-purpose intelligent technologies can help the disadvantaged area catch up in terms of technology. Fourth, AI as a Source of Positive Regional Ecological Spillovers. Spatial decomposition shows that 31.9% of the total policy dividend is in the form of indirect positive spillovers to interconnected regional areas. This suggests positive ecological spillovers rather than beggar-thy-neighbor enforcement displacement that has been observed in many older environmental regulations (Chagas et al., 2016). We posit that algorithmic architectures, open-source climate models (Rolnick et al., 2022), and unified environmental cloud platforms have low marginal replication costs, AI pilot zones generate positive regional ecological spillovers that enhance the ecological resilience of the entire urban agglomeration.

Summary

Integrating the Pressure-State-Response (PSR) model with the time-series global entropy weight method across a balanced panel of 280 Chinese cities from 2000 to 2022, this study employs a multi-method econometrics suite—including TWFE Multi-period DID, Double Machine Learning (DML), Bartik IV-2SLS, Coefficient-Based Causal Mediation, and Spatial Durbin DID—to investigate how National Next-Generation AI Innovation Pilot Zones reshape urban ecological resilience. Our main findings are: 1. AI pilot zone designation exerts a statistically significant, robust positive causal effect on composite urban ecological resilience (baseline TWFE estimate = 2.8500, DML = 2.8124,), which withstands heterogeneous treatment decompositions, 500-iteration Monte Carlo

permutation placebos, and augmented instrumental variable estimation, exhibiting persistent dynamic treatment effects over time;2. The PSR decomposition reveals an asymmetric structural profile: the policy dramatically suppresses pollution pressure (3.1200), steadily expands baseline biophysical carrying states (+1.4500), and aggressively enhances governance response (+4.2000);3. Green technology innovation (28.8% mediation share) and environmental enforcement intensity (25.6% mediation share) serve as the primary micro-level transmission conduits, reinforced by industrial structural upgrading (18.9% mediation share);4. Policy efficacy is amplified in old industrial bases, resource-dependent cities, and regions southeast of the Hu Line, while generating substantial positive spatial spillovers (indirect effects account for 31.9% of total spatial gains). Based on these findings, we propose three policy recommendations to maximize the ecological benefits of AI deployment: 1. Strengthen Enterprise Environmental Disclosure and Algorithmic Pollution Tracing: Municipal authorities should deploy IoT edge-computing and AI emission-tracing algorithms to construct holographic environmental credit profiles for major industrial polluters. Linking real-time emission monitoring with green credit scoring will incentivize enterprises to adopt clean production technologies and curtail pollution at the source;2. Deploy Targeted Technological Subsidies and Differentiated Regulatory Toolkits: Recognizing the higher remediation elasticity of old industrial bases and resource-dependent cities, central and provincial governments should establish dedicated intelligent green retrofit funds. Dynamically calibrating regulatory stringency and emissions-trading quotas to match local industrial structures and marginal abatement costs will prevent one-size-fits-all inefficiencies;3. Build Cross-Regional Computing Hubs and Environmental Data Sharing Networks: To dismantle administrative data silos, policymakers should establish inter-city cloud-computing hubs and shared ecological monitoring platforms. Leveraging the algorithmic spillovers of AI pilot zones will enable geographically connected cities to benefit from shared environmental intelligence, fostering coordinated cross-boundary ecological governance across urban agglomerations.

References

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. Journal of Economic Perspectives, 31(2), 3–32. https://doi.org/10.1257/jep.31.2.3

Bach, P., Chernozhukov, V., Kurz, M. S., & Spindler, M. (2022). DoubleML—An object-oriented implementation of double machine learning in Python. Journal of Machine Learning Research, 23(53), 1–6.

Ben-Michael, E., Feller, A., & Rothstein, J. (2021). The augmented synthetic control method. Journal of the American Statistical Association, 116(536), 1789–1803. https://doi.org/10.1080/01621459.2021.1929245

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-experimental shift-share research designs. The Review of Economic Studies, 89(1), 181–213. https://doi.org/10.1093/restud/rdab058

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. The Review of Economic Studies, 91(6), 3253–3285. https://doi.org/10.1093/restud/rdae007

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2024). Difference-in-differences with a continuous treatment. Journal of Econometrics, 240(2), 105670. https://doi.org/10.1016/j.jeconom.2023.105670

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 21(1), C1–C68. https://doi.org/10.1111/ectj.12097

Chagas, A. L. S., Azzoni, C. R., & Almeida, A. N. (2016). A spatial difference-in-differences analysis of the impact of sugarcane production on respiratory diseases. Regional Science and Urban Economics, 59, 24-36. https://doi.org/10.1016/j.regsciurbeco.2016.04.002

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Goralski, M. A., & Tan, T. K. (2020). The impacts of artificial intelligence on environmental sustainability and human well-being. Management Decision, 58(10), 2241–2259. https://doi.org/10.1108/MD-11-2019-1618

Gupta, P. K., Shree, V., Hiremath, L., & Rajendran, S. (2022). The use of modern technology in smart waste management and education: An overview. Environmental Science and Pollution Research, 29(48), 72471–72486. https://doi.org/10.1007/s11356-022-22793-9

Liu, L., Wang, Y., & Xu, Y. (2024). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. American Journal of Political Science, 68(1), 160–176. https://doi.org/10.1111/ajps.12723

Lyu, Y., Wang, W., Wu, Y., & Zhang, J. (2023). How does digital economy affect green total factor productivity? Empirical evidence from China. Science of The Total Environment, 857, 159428. https://doi.org/10.1016/j.scitotenv.2022.159428

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends in difference-in-differences. The Review of Economic Studies, 90(5), 2555–2591. https://doi.org/10.1093/restud/rdad018

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., ... Bengio, Y. (2022). Tackling climate change with machine learning. ACM Computing Surveys, 55(2), 1–96. https://doi.org/10.1145/3485128

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

Sant'Anna, P. H. C., & Zhao, J. (2020). Doubly robust difference-in-differences estimators for panel data. Journal of Econometrics, 219(1), 101–122. https://doi.org/10.1016/j.jeconom.2020.01.001

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, 225(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Wagner, K., Athey, S., & Imbens, G. W. (2022). Machine learning for causal policy evaluation in environmental economics. Resource and Energy Economics, 69, 101315. https://doi.org/10.1016/j.reseneeco.2022.101315

Zhang, Y., Wang, J., Liu, Y., & Zhao, J. (2024). The impact of the digital economy on urban ecological resilience: Empirical evidence from China's comprehensive big data pilot zone policy. Sustainability, 16(9), 3611.

Ziter, C. D., Pedersen, E. J., Kucharik, C. J., & Turner, M. G. (2019). Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. Proceedings of the National Academy of Sciences, 116(15), 7575–7580. https://doi.org/10.1073/pnas.1817561116

Appendix A: Policy Implementation Timing and Empirical Coding

City	Official approval date	Official year	Empirical treatment year
Beijing	2019-02-20	2019	2019
Shanghai	2019-05-22	2019	2019
Tianjin	2019-10-17	2019	2020*
Shenzhen	2019-10-17	2019	2020*
Hangzhou	2019-10-17	2019	2020*
Hefei	2019-10-17	2019	2020*
Deqing	2019-11-02	2019	2020*
Jinan	2020-01-23	2020	2020
Chengdu	2020-01-23	2020	2020
Xi'an	2020-01-23	2020	2020
Chongqing	2020-01-23	2020	2020
Guangzhou	2020-09-03	2020	2021*
Wuhan	2020-09-03	2020	2021*
Suzhou	2021-03-19	2021	2021
Changsha	2021-03-19	2021	2021
Zhengzhou	2021-11-13	2021	2022*
Shenyang	2021-11-13	2021	2022*
Harbin	2021-11-13	2021	2022*

* Note: For cities approved late in the calendar year (e.g., October-November), the empirical treatment year is coded as the following calendar year to accurately reflect the actual implementation and policy exposure duration. Similarly, Guangzhou and Wuhan (September 2020) are coded as 2021 based on the exact same rule (approval after June 30).

The Role of FinTech in Cross-Border Financial Risk Management: A Case Study of China Bohai Bank

Zhong Yang
Farabi International Business School, Al-Farabi Kazakh National University

Abstract

This article examines the relationship between digital infrastructure and AI-assisted cross-border risk monitoring through a documentary case study of China Bohai Bank. It compares the Bank's 2025 interim and annual disclosures, distinguishing reported facts, calculated changes and proposed mechanisms. The Hong Kong Branch's reported assets increased from HKD18.3 billion at mid-year to HKD21.689 billion at year-end, approximately 18.5%. Bank-wide technology investment was RMB1,274 million for the year; the disclosures record completed infrastructure projects alongside AI risk-analysis applications still in pilot phases. These findings support a distinction between infrastructure readiness and demonstrated AI effectiveness. The article develops three propositions: integration may account for gains attributed to AI; alert quality depends on review capacity and decision deadlines; and different risk horizons require different analytical methods. Proposed applications are matched to cross-border payment review, the Bank's own foreign-exchange exposure controls and onshore-offshore customer review. A three-condition evaluation design separates existing rules, integrated data with rules, and integrated data with rules and AI. Public evidence supports the institutional analysis but cannot establish detection accuracy or causal loss reduction. The contribution is a case-grounded explanation of where incremental AI value should be sought and how it could be tested.

Keywords: artificial intelligence; cross-border banking; financial risk; early warning; model governance; China Bohai Bank

1 Introduction

An ordinary trade payment can become difficult to assess when customer records, supporting documents and counterparty information sit in separate systems. A transaction amount may fall within an approved limit while the payment destination or beneficiary differs from the customer's established business. Conversely, an unusually large payment may reflect a legitimate new contract. The practical task is to bring these facts together before an operational decision is required.

Cross-border banking adds currency exposure and jurisdictional differences to this problem. Payment screening, customer credit assessment and the monitoring of market positions address related but distinct risks. Combining their information can improve context, but combining them into an undifferentiated risk score may obscure why an intervention is needed. An effective early-warning process must therefore specify both the risk being assessed and the action an alert is intended to support.

AI offers methods for finding patterns in large and heterogeneous datasets. The Financial Stability Board identifies applications in compliance, surveillance and fraud detection, alongside limitations arising from model opacity and dependence on data. These capabilities are relevant to banking, but they do not establish that an AI system will outperform existing controls in a particular institution.

China Bohai Bank provides a useful setting for examining this distinction. Its 2025 Annual Report records international settlement volume exceeding USD100 billion, 48% above the previous year, and qualification as a direct participant in the Cross-border Interbank Payment System (CIPS) (China Bohai Bank, 2026, p. 37). The report also describes integrated financial-market processing. These developments raise questions about how the resulting information could support risk monitoring; they do not, on their own, demonstrate improved detection.

This article asks what the Bank's disclosures establish about its cross-border and digital capabilities, how those capabilities could support an AI-assisted warning process, and what evidence would be needed to evaluate that process. The analysis retains the distinction between reported institutional developments, the author's interpretation and proposed operational arrangements throughout.

2 Literature and analytical approach

Risk data aggregation provides a starting point for examining AI-assisted monitoring. The Basel Committee on Banking Supervision links weaknesses in aggregation and reporting to banks' inability to identify exposures and concentrations promptly. Its principles address governance, infrastructure and the quality of risk information. This article uses those principles as an analytical reference, without asserting that every provision applies directly to the case institution.

AI research extends the discussion from assembling information to interpreting it. Aldasoro et al. examine how successive forms of AI affect financial intermediation, insurance, asset management and payments. They identify improvements in information processing alongside challenges involving privacy, discrimination and interconnectedness. Their emphasis on accountability and human oversight is relevant to an alerting process in which staff must explain and justify decisions.

Rules and statistical models serve different purposes within that process. A rule can express an explicit internal limit or a defined screening condition. An anomaly model can indicate that behaviour differs from a reference pattern, even when no individual variable exceeds a fixed threshold. An anomaly is not proof of misconduct, and a model score is not a probability of wrongdoing unless it has been appropriately defined and calibrated. These distinctions affect both the language of alerts and the authority given to the system.

The analytical issue is attribution: when data platforms, operational processes and AI applications change together, which change produces a better risk decision? A generic sequence from data collection to human review does not resolve that issue. This article instead examines the sequence of disclosed capability changes and the control tasks they could affect. It develops case-specific propositions about incremental AI value, which remain subject to empirical testing rather than being presented as established general laws.

3 Research design and evidence

3.1 Case selection and source boundaries

China Bohai Bank is selected because its 2025 disclosures describe several changes within one institution: expanding international settlement activity, development of its Hong Kong Branch, new data and transaction infrastructure, and pilot AI applications. Their coexistence makes the case suitable for examining the boundary between digital integration and AI capability. It does not make the Bank representative of all commercial banks, nor does the study select it on the basis of demonstrated AI success.

The evidence consists of selected passages from the 2025 Interim Report and 2025 Annual Report, supplemented by the Hong Kong Monetary Authority's register for licensed status. The observation points are 30 June and 31 December 2025; annual flows refer to the full reporting year. The case material is concentrated in the interim report's information technology and overseas branch sections, and the annual report's cross-border business, financial-market,

technology, overseas branch and risk-management sections. Page references below use printed report pagination.

3.2 Evidence extraction and comparison

Each selected observation is classified as a reported fact, an author calculation or an analytical inference. Numerical observations retain their reporting period, currency and institutional scope. Project descriptions are distinguished as construction, launch or pilot activity. A bank-wide system launch is not coded as deployment of a cross-border AI model unless the disclosure establishes that application. The qualitative evidence register in Section 4 makes these classification decisions visible.

The comparisons use two disclosed observations within 2025. Percentage change is calculated as the year-end value divided by the mid-year value, less one. Full-year technology investment minus first-half investment yields an implied second-half amount, conditional on consistent reporting scope. This is a descriptive comparison, not a before-and-after estimate of a treatment effect. Annual settlement growth, branch asset growth and technology staffing describe different populations and are not combined into a productivity ratio.

3.3 Reliability and limits

The two reports provide temporal corroboration within the same reporting institution, not independent verification of performance. The HKMA register corroborates licensed status only. No interviews, transaction-level observations or independent coding exercise were conducted. The selection of relevant passages is transparent but is not an exhaustive systematic review of all Bank disclosures. An unreported outcome is classified as unavailable, not as evidence that the capability or outcome does not exist.

The study can establish what changed in the disclosed institutional setting and develop explanations consistent with those observations. It cannot measure missed events, model discrimination, staff workload or avoided losses. Sections 6 and 7 consequently distinguish the proposed applications and evaluation protocol from the descriptive results reported next.

4 Descriptive results and interpretation

4.1 Public indicators and their measurement scope

The 2025 Annual Report records international settlement volume above USD100 billion and a year-on-year increase of 48%, together with CIPS direct-participant qualification (China Bohai Bank, 2026, p. 37). These observations establish a growing business and a change in payment connectivity. Settlement value does not reveal transaction count, message-processing latency or the number of cases requiring investigation. An exact prior-year settlement value cannot be recovered from a rounded growth rate and a reported lower bound.

Table 1 Public indicators at the two observation points

Indicator	Mid-year 2025	Year-end 2025	Author calculation
Technology investment, RMB million	518 January–June	1,274 January–December	Implied July–December: 756
Technology personnel, persons	1,497	1,473	−24; approximately −1.6%
Hong Kong Branch assets, HKD billion equivalent	18.3	21.689	+3.389; approximately +18.5%

Sources: China Bohai Bank (2025), pp. 36–37; China Bohai Bank (2026), pp. 45–46. Investment is a cumulative flow; personnel and assets are point-in-time stocks. Technology measures are

bank-wide, not cross-border AI measures. Calculations assume comparable reporting scope; branch figures reflect reported HKD equivalents.

The Hong Kong Branch's reported assets rose by HKD3.389 billion between the two observation points, approximately 18.5%. Because the mid-year figure is rounded and both figures are reported in HKD equivalents, the change should be interpreted approximately. It is not a constant-currency estimate or a measure of risk exposure after hedging. Nevertheless, it establishes that the overseas branch's balance-sheet scale changed materially during the period, making cross-entity exposure and customer-information coordination a relevant case question.

On a consistent reporting basis, implied second-half technology investment was RMB756 million, calculated as RMB1,274 million less RMB518 million. The year-end technology personnel count was 24 below the mid-year count, approximately 1.6%. Investment flows and staffing stocks are different measures; these figures do not establish higher productivity, staff reductions in risk review or expenditure on AI specifically. They do show why spending growth cannot substitute for evidence about operational review capacity.

4.2 An observable distinction between launch and pilot activity

The interim report describes ongoing construction of core, funds-management and data infrastructure and preliminary design and evaluation of an intelligent-agent platform (China Bohai Bank, 2025, p. 36). At year-end, the annual report describes launches of the distributed core system, the main funds-management project and core data-lake functions. In the same section, AI risk analysis and automated corporate due-diligence report generation are described as pilot scenarios (China Bohai Bank, 2026, p. 45). The disclosed sequence therefore distinguishes implemented infrastructure from applications whose performance is not established by the report.

Table 2 Selected disclosure evidence and the limits of inference

Observation	Reported evidence	Interpretive boundary
Cross-border activity	Settlement value above USD100 billion; growth of 48%; CIPS qualification. Annual report, p. 37.	Business scale and connectivity; not transaction count or detection quality.
Digital infrastructure	Construction at mid-year; launches reported at year-end. Interim report, p. 36; annual report, pp. 40, 45.	Infrastructure availability; coverage of each cross-border population remains to be verified.
AI applications	AI risk analysis and corporate due-diligence report generation described as pilots. Annual report, p. 45.	Evidence of pilot activity; no disclosed comparative performance for the proposed uses.
Existing exposure controls	Exposure analysis, limits, hedging and stress testing. Annual report, p. 52.	An existing baseline against which added AI value must be tested.

Source: author classification of selected passages in China Bohai Bank (2025, 2026). The register is documentary evidence, not an independent audit or a score of the Bank's actual capability.

The strongest supported finding is this difference in disclosed development stages. Integration and AI are changing within the same institution, but the reports offer stronger evidence about systems becoming available than about AI improving risk decisions. It would

therefore be premature to interpret the case as proof that AI has reduced cross-border risk. A more defensible inference is that integration creates a testable setting for assessing incremental AI value.

4.3 Existing controls constrain the proposed contribution

The Bank already reports currency exposure analysis, limits, derivatives used to manage foreign-exchange exposure and regular stress testing. Its disclosed cumulative foreign-exchange exposure ratio within the consolidated scope was 1.26% in 2025 (China Bohai Bank, 2026, p. 52). This aggregate indicator cannot establish risk for a particular currency, branch or transaction, but the control description shows that a new proposal must build on existing exposure management rather than assume its absence.

Similarly, intraday liquidity management is described in terms of per-currency funding requirements, while daily liquidity alert monitoring and periodic stress tests are also reported (China Bohai Bank, 2026, pp. 49–50). These are distinct decision horizons. Payment-level anomalies, intraday funding requirements and portfolio stress cannot usefully be treated as a single prediction task. This distinction informs the proposed applications below.

5 Proposed risk identification and early-warning process

5.1 Establishing a usable transaction record

The process would begin by linking customer identifiers, payment instructions, counterparty records and relevant exposure information. A risk officer should be able to trace the information in an alert to its source and determine when it was last updated. Inconsistent names, duplicate customer records and missing timestamps should be treated as data-quality issues before they are interpreted as unusual behaviour.

Integration also requires limits on what information is available to whom. The proposed design would give staff access according to their role and the applicable arrangements for information sharing. Where records cannot be combined, the system should make that limitation visible. It should not silently treat unavailable information as evidence of low risk.

5.2 Combining rules with anomaly detection

Existing mandatory screening and approved limits would remain explicit controls. Statistical or machine-learning models could supplement them by comparing a payment with the customer's previous activity or an appropriate peer group. Candidate indicators include changes in payment size, frequency, destination and counterparty relationships. Their selection would depend on the risk objective and available data, rather than on the assumption that more variables necessarily produce a better model.

Different customers require different reference patterns. Seasonal trade, a new business line or a newly onboarded customer can make a historical comparison misleading. For customers with little history, the process would need a documented fallback, such as closer documentary review or a carefully selected peer comparison. The proposed system should report why an observation appears unusual and identify important data gaps.

5.3 Assessing the signal

The assessment stage would distinguish a possible compliance issue from market exposure, counterparty deterioration or an operational exception. Those categories may involve different evidence, urgency and responsible teams. A single aggregate score should not conceal a mandatory restriction or allow a weak signal in one category to cancel a serious signal in another.

Thresholds would need to reflect the consequences of missed events and unnecessary interventions. They cannot be specified credibly from public reports alone. Any pilot should examine how alternative thresholds affect both the quality of alerts and the number of cases that staff can investigate. Changes in transaction patterns would prompt a review of model performance rather than automatic, unapproved changes to the model.

5.4 Routing warnings to an accountable reviewer

An alert would contain the relevant transaction, the triggering rule or model output, the principal contributing factors and the required review route. Lower-priority anomalies might be monitored or placed in a review queue. Signals requiring intervention before settlement would need a defined decision deadline and an authorized reviewer. The categories would be operational priorities, not findings that a customer had acted improperly.

Review capacity is part of the design. A system that creates more warnings than staff can assess may lengthen queues while leaving urgent cases unresolved. Alert ownership, escalation for overdue cases and procedures for system outages should therefore be specified alongside detection logic. Payment holds or other interventions would follow the Bank's authorized procedures and applicable requirements, not a generic model score alone.

5.5 Recording decisions and controlling feedback

The reviewer would document the evidence considered, the decision and any unresolved questions. A payment supported by a new commercial contract might be cleared even though it differs sharply from historical behaviour. An alert with incomplete documentation might require further investigation without yet being classified as a confirmed risk event.

These distinctions matter for subsequent model development. Closing an alert is not equivalent to proving that no risk existed. Feedback records should distinguish confirmed events, supported legitimate activity and unresolved cases. Any retraining would require quality checks, validation and approval. Investigation outcomes would inform model improvement through a controlled process rather than being incorporated automatically as unquestioned labels.

6 Case adaptation and proposed explanatory propositions

6.1 Cross-border payment review and the contribution of integration

The case anchor is the Bank's disclosed international settlement growth and CIPS qualification, rather than an assumed absence of payment controls. A proposed pilot would cover a defined subset of corporate cross-border RMB payments, conditional on access to payment records and supporting customer information. The first task would be to verify that the relevant payment-processing system can link those records; CIPS qualification alone does not demonstrate this internal linkage or specify a review deadline.

The proposed comparison would first add linked customer history and counterparty information to existing rules. Only after measuring that change would it add an anomaly model. An apparent improvement over the original process may result from fewer missing records or better entity matching, even when the detection logic remains unchanged. The funds-management system's disclosed scope is financial-market business, so coverage of this corporate payment population must be verified separately before selecting it as the technical platform.

Proposition 1 is that a comparison against fragmented-data rules will overstate AI's incremental contribution when data integration itself improves the decision. Evidence consistent with this proposition would be a substantial gain from integration with rules, followed by a smaller or absent gain from adding AI. It would be challenged if integration produced little improvement while the AI condition produced a reproducible gain under the same data, population and review capacity. This is a proposed empirical distinction, not a result observed in the public reports.

6.2 Bank foreign-exchange exposure and the importance of decision horizons

The relevant exposure is the Bank's own position, rather than the foreign-currency obligation of an unrelated hypothetical importer. The disclosed baseline already includes exposure analysis, limits, hedging and stress testing. A suitable extension would reconcile position, settlement and hedge records by currency and identify stale, missing or inconsistent inputs before those inputs affect established calculations. An anomaly in a feed could justify operational investigation without being treated as a forecast of exchange-rate movements.

The financial-market system described in the annual report offers a closer institutional match for this task than for all corporate payments. The Assets and Liabilities Management Department is identified as leading relevant market-risk mechanisms, providing a disclosed organizational anchor for the proposed review route (China Bohai Bank, 2026, pp. 51–52). Access to the necessary records and the proposed workflow remain implementation requirements, not reported facts.

Proposition 2 is that incremental AI value differs by risk horizon: document or feed-quality analysis may improve the inputs to exposure management without outperforming deterministic calculations of exposure. The proposition would receive support if corrected data reduced reconciliation exceptions or stale-position decisions while an AI exposure predictor added little beyond the existing calculations. It would be challenged by a validated predictive model that improved the chosen decision outcome using the same timely inputs. The aggregate 1.26% exposure ratio cannot adjudicate between these explanations.

6.3 Onshore and offshore customer review under capacity constraints

The Hong Kong Branch's disclosed wholesale and financial-market activities, including cross-border services and trade financing, make shared-customer review a closer case match than generic retail fraud detection (China Bohai Bank, 2025, p. 37; 2026, p. 46). The annual report's pilot work on corporate due-diligence reports suggests an application near this task, although it does not establish that the pilot is used in Hong Kong or shares data across jurisdictions.

The proposed application would compare authorized customer records and supporting documents for inconsistencies in entity identity, transaction purpose or financing relationships. Any generated summary would preserve links to the underlying passages so that reviewers could check it. Information-sharing permissions, source availability and responsible staff would need to be established before a pilot. The model would assist discrepancy review, not make an autonomous legal or compliance determination.

Proposition 3 is that an alerting system can improve detection metrics yet fail operationally if the resulting reviews cannot be completed within the decision window. Let N denote eligible cases in a defined interval, a the fraction routed to review, h average review minutes per case and K available reviewer minutes. For a stable queue in that interval, the proposed process requires $N \times a \times h$ to remain below K, with headroom for uneven arrivals and urgent cases. This accounting relationship is a design constraint, not an estimated model of the Bank's workload. None of these parameters is disclosed in the case sources.

The proposition would receive support if increased alert volume improved case identification but raised overdue reviews or delayed legitimate payments. It would be challenged if additional alerts were processed on time through shorter reviews or added capacity without sacrificing quality. Mandatory screening requirements would remain in force regardless of workload; a capacity breach would trigger staffing or process escalation, not silent suppression of required controls.

6.4 Contribution and scope of the propositions

The article's contribution is the connection between an observed development sequence and these distinct tests of incremental value. Data integration is a competing explanation for better detection; risk horizon determines the appropriate method; and review capacity mediates whether detection becomes timely action. Together, they explain why a bank could report major technology investment and successful system launches while the effectiveness of AI remained unresolved. The contribution is analytical and case-grounded, not a claim that these general concepts have never appeared in prior research.

7 An evaluation protocol that separates competing explanations

7.1 Three comparison conditions

A subsequent study should compare three conditions on the same eligible transaction population: A, the existing data and rule-based process; B, integrated and reconciled data with the same rules; and C, the same integrated data and rules supplemented by the proposed AI component. The change from A to B estimates the contribution associated with integration in the evaluation setting; the change from B to C isolates the incremental AI component more closely than a simple before-and-after comparison. These contrasts would be descriptive in a retrospective replay and would require stronger assignment and measurement conditions for a causal interpretation.

For the payment pilot, review budgets and mandatory controls would be held constant across conditions. For the exposure pilot, the baseline would retain approved deterministic calculations; candidate AI would be evaluated on a defined added task such as detecting feed exceptions. For the customer-document pilot, assisted and unassisted reviews would use comparable cases and independently assessed reference outcomes. The three applications require separate endpoints rather than a pooled accuracy score.

7.2 Outcomes and reference decisions

For payment review, the main measures would be confirmed relevant findings per completed review, known-event detection, time to decision and legitimate-payment delay. For exposure monitoring, measures would include verified reconciliation exceptions, stale inputs corrected before the control deadline and false operational alerts. For document review, measures would include verified discrepancies found, unsupported statements in generated summaries and reviewer time. These are proposed measures; the present article reports no observed values for them.

Reference decisions would need a defined event taxonomy and independent case review where feasible. A case closed without further action would not automatically become a confirmed negative label. Sampling unflagged cases would help assess missed events, although inference to the full population would require appropriate sampling weights and recognition of incomplete investigation. Precision among reviewed alerts would not be presented as recall for all risk events.

7.3 Timing and analysis

Development data would precede a locked evaluation period. All features would be restricted to information available at the original decision time, while outcome labels could use a clearly defined later follow-up window. Repeated transactions from the same customer create dependence, so uncertainty estimates should account for customer-level clustering rather than treating each transaction as independent. Results would be stratified by currency, customer-history availability and payment route where sample size permits.

The study would compare B with C using a prespecified outcome and review budget. A useful result would require an improvement large enough to matter operationally, with uncertainty reported, without unacceptable deterioration in delays or missed events. Minimum sample size and acceptance thresholds cannot be specified from these public disclosures: they depend on event prevalence, effect size, dependence among cases and the Bank's approved tolerances.

7.4 Deployment and alternative explanations

A retrospective replay would test detection logic but could not establish how staff respond in live operations. A shadow phase could then measure arrivals, review time and explanation quality without allowing the experimental system to replace mandatory controls. Any controlled live study would require approved arrangements, with the same baseline protections maintained across conditions.

Observed differences might reflect changing customers, currency conditions, investigator experience or data-quality improvements rather than AI. The three-condition comparison addresses one important source of confounding but does not remove all of them. Matching the calendar period and population, documenting changes in staffing and checking the completeness of outcomes would strengthen inference. Only after such evaluation could the Bank credibly claim an operational improvement attributable to the proposed AI component.

8 Conclusion

The descriptive evidence identifies a specific institutional configuration at China Bohai Bank: cross-border business expansion, a larger reported Hong Kong Branch balance sheet, completed digital infrastructure projects and AI applications described as pilots. These developments make the Bank a useful case for examining incremental AI value, but provide no direct estimate of improved detection or reduced losses.

The analysis advances beyond a generic early-warning workflow by separating three questions. Did integration improve the information used by existing controls? Did AI improve a defined task beyond that integrated baseline? Could staff act on its outputs within the relevant deadline? The proposed applications match these questions to the Bank's disclosed payment activity, financial-market exposure controls and onshore-offshore corporate business.

The public-data comparison supports the institutional interpretation; the explanatory propositions and evaluation protocol remain to be tested. Access to anonymized cases, system timestamps, documented investigation outcomes and review-capacity information would permit that next stage. Until then, the paper's defensible contribution is a transparent documentary case and a falsifiable account of where additional AI value may arise.

References

Aldasoro, I., Gambacorta, L., Korinek, A., Shreeti, V., & Stein, M. (2024). Intelligent financial system: How AI is transforming finance. BIS Working Papers, No. 1194. Bank for International Settlements.

Basel Committee on Banking Supervision. (2013). Principles for effective risk data aggregation and risk reporting. Bank for International Settlements.

China Bohai Bank Co., Ltd. (2026). Annual report 2025.

Hong Kong Monetary Authority. (2026). China Bohai Bank Co., Ltd. Register of Authorized Institutions and Local Representative Offices. Retrieved September 12, 2026, from

Smith, A. (2019). AI and risk management in global banking. London: Palgrave Macmillan.

BoHai Bank. (2022). Annual Report 2021: Embracing digital transformation. Tianjin: BoHai Bank.

People's Bank of China. (2022). Report on the Development of the RMB Cross-border Payment System (CIPS) (2021). Beijing: People's Bank of China.

From Screening to Retention: Extending AI Bot Roles Across the Employee Lifecycle: A Case Study of a Mobile Game Studio

Milena Kazimir
Bachelor's degree in Political Science and World Politics; Senior Recruiter, IT

Annotation:
AI bots in HR are generally associated with recruitment and are rarely carried over to other stages of the ELC. Moreover, while traditional bigtech companies are often the subject of the research, the gamedev industry - with its project-based structure, heterogeneous teams, and crunch culture - remains an under-researched case in the context of AI and HR. On the basis of this context, our research question was: How does AI adoption vary across the stages of the employee lifecycle? Focusing empirically on recruitment as a best-case scenario, to what extent does this pattern hold across heterogeneous roles within a game development studio? As a method, the insider single case-study on the example of an anonymised small game studio was chosen. We conducted the comparison of HR metrics before and after implementation of AI tools on the team level and, separately, creative roles. The key finding was that process metrics reduced by 60–77% in both segments; at the same time, those in creative roles achieved comparable or better results than the team as a whole, contrary to the expectations of the heterogeneity hypothesis. Our contribution consists of the development of a theoretical framework (Differential AI Adoption), a refinement showing that tool design (custom-built vs. generic) may matter more than role type; and a practical recommendation for small studios.

keywords: AI bots, HR, Employee Lifecycle, Recruitment, gamedev.

Introduction

In recent decades, the gamedev industry has taken a confident place in the global economic market, earning $188.8 billion in 2025, which is more than the cinema and music industries all together (industry analysis: IconEra, 2025). However, the industry is experiencing two contradictory trends: on the one hand, there have been massive layoffs, with around 15,600 workers getting fired in 2024 and just above 9,100 in 2025 (industry analysis: GamesIndustry.biz, 2025); on the other hand, the talent shortage persists, resulting in a 40% increase in job openings in recent years (Bain&Co, 2023). The industry itself is not broad, consisting of nearly 1 million people all around the world (WifiTalents, 2026), which makes it difficult to find highly-skilled candidates for some demanding roles such as Game Designers or Engine Developers, leaving vacancies open for months and pushing companies to change their approach and hire specialists from other industries. Another disturbing factor is the short periods of work of employees: according to the industry analysis by Bain & Co. (2023), the game industry has a 10% attrition rate, while for tech companies, it is only 6%.

Both the necessity to tackle the low attrition rate figures and an influx of specialists outside the game development bubble have been changing the way how game companies operate inside. For decades, the gamedev industry has been building its operational processes on the inspiration and passion of its employees for games, preserving the chaotic startup approaches, culture of

crunches, and poor compensation. Recent changes in the industry push companies to implement new practices: work-life balance, decent salaries, career development assistance, ESG commitments, well-established workflows, and benefit systems (Bain&Co, 2023). The implementation of modern operational processes falls on the shoulders of HR and project managers, who look for the optimal methods and instruments, including the introduction of AI agents.

AI agents have been widely introduced in recruiting cycles, in particular, screening resumes, writing job posts, sourcing specialists, transcribing interviews, and asking preliminary clarifying questions from candidates via AI bots. According to the World Economic Forum (2025), 88% of companies worldwide use AI in some respects during the initial stages of screening candidates. As a result of this wide use, AI agents are dominantly seen as the tool for recruiting and are rarely integrated in other spheres of HRD. It pushed us to propose *research questions*: How does AI adoption vary across the stages of the employee lifecycle? Focusing empirically on recruitment as a best-case scenario, to what extent does this pattern hold across heterogeneous roles within a game development studio?

We decided to build our case study in the context of the gamedev industry for several reasons. First of all, because of its project-based structure of work: short development cycles of 1-3 years create predictable release-related employee attrition figures. Secondly, crunch culture is a systematic attrition factor in the industry (Bain & Co., 2023), which allows for monitoring involvement in real time. Thirdly, the heterogeneity of teams, from highly specialised programmers to creative roles (artists, sound designers, etc.) with different motivation factors, makes standardised AI retention risk models, learned on IT-tech datasets, of limited capability and proves the necessity of a case study on a game development sample.

The research article will first present a literature review broken down by stages of the employee life cycle, then develop a theoretical framework, discuss a methodology, and analyse the empirical data and their alignment with the literature.

1. Literature Review

1.1. AI and HR-Tech: General Review

The use of AI tools in HR processes has experienced a boom in recent years. The idea of digital transformation in HR first appeared in academic literature in 2015. However, research on the topic remained relatively limited until 2018, when AI adoption among organisations increased from 20% to 47% (Bindra et al., 2025). Since then, the scope of analysis has shifted from basic automation to the complex forms of relationships between AI and human beings (Bindra et al., 2025).
It is worth noting that, unlike some other IT domains, such as coding, the HR sphere has not yet fully adapted to the integration of AI into its processes. According to the survey conducted by McKinsey in 2025, 88% of industrial organizations claim about the regular use of AI in at least one of the business functions. Nevertheless, scaling of AI-agentic solutions accounts for only 23% of cases, and in no single function does the percentage of such organisations exceed 10%, which is relevant for the HR Department as well.
Tambe et al. (2019) were the first who conducted a systematic analysis of this gap, identifying four barriers that prevent the use of data science in HR: 1. HR indicators such as engagement and culture fit do not translate well into statistical data and require contextual and textual analysis on a case-by-case basis; 2. as a tendency, there are small samples inside companies, which makes the

generalization more complicated and introduces the lack of statistical data; 3. when it comes to decisions about human beings, more transparency is required; 4. employees tend to stand against the algorithm-based management decisions. Even nowadays, these obstacles have not lost their relevance.

Despite that, in recent years, there has been progress in successfully integrating AI into HR processes. Baranyi (2025) analysed 94 Scopus-indexed studies and proposed a three-level typology - Manual (without usage of AI), Digitalised (partly automatised processes, preserving the 'human-in-the-loop' practice, when the technology serves as an assistant and does not make any decisions) and Digitally Transformed (AI actively reshapes the core of the function and operates autonomously) Practices - to describe the depth of AI integration into recruitment processes. She claims that more and more functions are now digitised or even digitally transformed, meaning the rise in popularity of AI in HR. Ekuma (2024) identifies a similar change in employee training and development, showing that AI is increasingly used not only for routine tasks but also in broader HR decision-making and strategy.

1.2. Screening and Recruitment

AI-based technologies are increasingly used in the early stages of recruitment. According to reported data, 87% of organisations use these tools as part of their hiring processes. In financial terms, the recruiting sector was worth approximately $704.54 million in 2025 and is projected to exceed $1.1 billion by 2032, growing at an annual compound rate of 6.8% (Kumar, 2026). The major driving factor for introducing AI instruments in recruitment is "savings in both costs and time, minimising risks, enhancing productivity, and increasing certainty in decision-making" (Köchling and Wehner, 2020, 796).

AI tools can be used for several different types of recruitment activities: publishing job ads, searching for candidates, handling applications, selection, assessment, and process coordination (Black & Van Esch, 2020). During the initial stage, job advertisement, recruiters can use AI to prepare job postings and adapt them to different platforms targeting different categories of job seekers (Chen, 2023). Not only that, AI may analyse the profiles of candidates and reach out to them on their own with standard messages linking their preferences and experience with vacancies (Black & Van Esch, 2020).

When looking for candidates, recruiters source them on different platforms using special Boolean requests. Research illustrates that AI-based sourcing algorithms outperform traditional tools such as LinkedIn Recruiter in terms of finding the most suitable candidate profiles (Slyakovskiy et al., 2025). What is more, AI may be used for scraping candidates' CVs on job platforms or their personal pages on LinkedIn to form a base of potential candidates for outreach (Black & Van Esch, 2020).

During the phase of selection, AI assess the candidate's resume, identifies the keywords in their work experience and matches them against the job requirements. Based on this assessment, the algorithm may automatically decline not suitable candidates and rank the remaining ones in order of priority and suitability (Leong, 2018). Actually, this particular integration (which is also called ATS, or Application Tracking System) is used widely across the globe, since it greatly reduces the time of recruiters who are required to process 400–500 applications for 5 or more vacancies simultaneously. Research shows that AI tools outperform humans in screening candidates by at least 25% (Chen, 2023). Moreover, even though we could not find academic sources with large datasets, Sharma analysed the case of integrating ATS in the screening process at L'Oreal and found that it reduced time to review resumes from 40 minutes to just 4 (Black & Van Esch, 2020).

After the screening conducted by a recruiter, AI participation may take various forms. Black & Van Esch (2020) offer gamified tests that can assess a candidate's abilities in greater depth and video-recorded interviews with AI-enabled technologies that ask candidates questions and request them to upload recorded answers. After that, algorithms analyse their responses, mimics and tone of

voice, match them with what average and best-performing candidates looked like in the past, and make the verdict.

Finally, AI may support the coordination between a recruiter and a candidate throughout the whole process. Instead of filling in the form by hand, an applicant may write such information as salary expectations, location, technology stack, years of experience, etc. to the chatbot, which will automatically enter the required information in the questionnaire (Chen, 2023). Furthermore, it is likely that AI bots may send notifications about the upcoming interviews, generate automatic rejections or schedule the meetings.

Overall, all these instruments reduce time-to-hire, meaning "the average number of days from the posting of a job vacancy to the formal hiring of a candidate" (Priyadarsini & Sreejith, 2026, 6), and administrative overload. Existing studies show the positive correlation between the integration of AI in recruitment and the reduction of this metric (Chandad et al., 2026; El Ouakili, 2025; Upadhyay & Khandelwal, 2018).

However, the widespread adoption of AI tools in recruiting raises critiques and controversial biases. To illustrate, Köchling and Wehner (2020) conducted a review which shows that algorithmic decision-making in recruitment systematically results in discrimination, because the algorithm itself is trained on historical data containing inequality. Other studies have reached similar conclusions, suggesting that historical employment data may contain gender, racial, and socio-economic biases that AI recruitment systems can reproduce in hiring decisions (El-Ghoul et al., 2024; Rigotti & Fosch-Villaronga, 2024; Sheard, 2025).

To prove the point, an experimental study by Ip (2025) demonstrates that gender bias in recruitment algorithms can significantly limit women's access to high-paying positions in traditionally male-dominated professional fields. In contrast, if the algorithm is manually retrained on a new, more up-to-date sample, representation of women among candidates at various stages of the selection process increases.

This opacity - a so-called problem of the "black box" - poses serious ethical challenges: if a candidate does not understand why the system has declined them, and a recruiter is not able to explain the logic behind the algorithmic decision, it becomes nearly impossible to tackle the discriminatory pattern.

In order to tackle this issue, the EU has proposed the EU AI Act. It establishes a risk-based classification framework under which AI systems listed in Annex III are generally considered high-risk (European Parliament & Council of the European Union, 2024, Art. 6(2)), meaning they should be carefully observed and controlled. Some of the demands are already in use: since February 2025, it is prohibited to use emotion recognition tools during interviews, and rules on AI literacy for staff have been introduced. The core set of requirements (mandatory human supervision, testing for bias using the company's own data, logging decisions) should be enacted later.

1.3. Onboarding

According to Wandhe (2026), high-quality and personalised onboarding programs increase retention rates by up to 82% and productivity by 70%. To effectively serve its purpose, the process of onboarding should start from the first day of new workers and continue throughout their first year. It should "consist of standard, organisation-specific components but also should include customisation to meet the needs of the individual, accounting for their personality and the experiences they bring with them as they join the organization" (Blount, 2022, 57). The process itself is demanding in terms of time and resources invested by HR and project managers, one possible solution of which is the introduction of AI into the flow.

Generally, three types of AI-assisted instruments integrated for onboarding purposes may be identified: AI assistants, typically incorporated into HRIS; AI chatbots; personalised training systems (Sudhakar, 2022; Wandhe, 2026; Xiang et al., 2023).

AI assistants may provide support to new employees by sending notifications, guiding them through the structure of a company, making notes or even matching them with colleagues with similar duties or interests, involving them in knowledge exchange and collaboration (Jasiak-Kaczmarek, 2025; Ritz et al., 2023).

As for training systems, AI can provide additional guidance during the learning process (Krishnan et al., 2024; Nawaz et al., 2024). AI tools can help identify training needs and adapt learning materials to individual employees based on their skills and preferences. Some systems are also able to provide virtual mentoring (Jasiak-Kaczmarek, 2025).

Finally, AI chatbots have the most universal use: instead of going to HR specialists, new employees may ask bots about benefits, values, policies, procedures, and documents (Wandhe, 2026). In contrast to traditional guidelines and FAQ checklists, bots support active conversations with employees, provide personalised answers, and actively encourage questions, thereby facilitating the time needed to get acquainted with a new workplace and shortening the adaptation period (Jasiak-Kaczmarek, 2025). We were not able to find solid academic researches, but several industry sources (for example, the Brandon Hall Group and Gallup; unfortunately, the exact figures from the report were not publicly available) agree that personalisation and immediate response instead of hours waiting reduce the time-to-productivity, which is a quantitative metric measuring the time that a new employee needs to reach key performance indicators and start contributing to a company (Society for Human Resource Management [SHRM], 2024).

However, chatbots do not always understand user requests correctly. This can be a problem when questions are complex or involve emotions or organisation-specific terminology (Wandhe, 2026). Employees may become frustrated if they receive an incorrect answer and even get angry or escalate it to the HRD, which will increase the time managers spend on the newcomer, rather than reducing it as originally intended. Moreover, Kumar and Singh (2024) postulated that the use of AI tools should be approached with caution, since it is the duty of organisations to ensure data security and to assure their employees that their private conversations remain confidential.

1.4. Engagement & Performance

The continuous evaluation of engagement and performance of the employees is crucial for managers: it allows them to allocate staff to tasks appropriately, to identify problem areas promptly, and to predict retention and attrition rates in the company. Generally, there are two broad ways of evaluating employee involvement and productivity: continuous feedback and annual performance reviews.

Performance reviews are formal meetings in which employees and managers discuss work results, recognise achievements, and identify areas for development, usually based on the employee's performance over the previous year (Maganti, 2025). For decades, this evaluating instrument has been widely used across all sectors of the economy. However, nowadays, it is claimed by many to be outdated: the criticism is primarily directed at annual performance reviews' inability to influence changes in real time, delayed feedback, and their failure to keep pace with the rapid changes in business processes that are now commonplace (Viterouli et al., 2024). As a response to static, rare evaluations, continuous feedback has been introduced.

Continuous performance feedback means the regular exchange of feedback between managers and employees throughout the year. Rather than relying on a single annual review, organisations may use weekly or monthly check-ins to discuss progress, performance, and areas for improvement (Traynor et al., 2021). According to the research by Maganti (2025), this type of feedback is more welcomed by employees (76% preferability in contrast with 24% for annual reviews) and is found more effective.

What is interesting is that it affects the rate of employee involvement in return: there is a positive ($r = 0.81$) correlation between the frequency of feedback and employee engagement (Maganti, 2025, 1046). Several other studies (Giamos et al., 2023; Idrus, 2025; Rivera et al., 2021) prove the

point and discover that regular feedback can enhance commitment and engagement by providing employees with an understanding of their progress, enabling real-time work course correction, and remaining aligned with organisational goals and expectations.

However, there are several limitations, the most important of which is the time required from managers to deliver feedback permanently (Idrus, 2025). To deal with this problem, AI-based solutions may be integrated into pulse surveys and sentiment analysis. AI algorithms may assess digital sources and footprints and conduct pattern analysis: an increase in the response time for emails, the ratio between off-work emails and working-hour emails, a reduction in participation in non-essential discussions, a change in tone in emails and corporate messengers, - which may reveal early stages of disengagement and low performance (Ch & Singha, 2024; Gelbard et al., 2018). Pulse surveys are generally conducted in the form of short questionnaires pushed by AI bots, which may also be then evaluated through sentiment analysis. The problem with the methods mentioned above is an ethical one. While analysis of open corporate chats in Slack or Teams raises no issues, access to private correspondence could undermine staff confidence in the company and run counter to the EU AI Act.

While pulse surveys and sentiment analysis generally present qualitative data and rarely make any numerical findings, the time-to-productivity metric mentioned above serves as a quantitative tool. According to SHRM (2024), it is not an isolated onboarding metric, but a part of a more extended system of performance indicators, which may be easily integrated into the AI dashboard next to the engagement score. Continuous feedback allows monitoring time-to-productivity dynamically, week-over-week, instead of a single measurement at the end of the trial period.

1.5. Retention and Prediction of Attrition

Attrition is defined as a voluntary or involuntary departure of an employee from a company (Hinge et al., 2023). Since it affects the efficiency of a company in terms of its productivity and stability and leads to time and money costs, the metrics of retention are scruitinisly monitored by managers. Generally, there may be found some exact numerical figures such as the age, length of service within the company, internal salary and its correlation with the average salary for a particular specialist in the market, performance KPIs, etc., based on which Machine Learning models may be developed to predict the possibility of turnover. For instance, Nandal et al. (2024) proved that ML models can predict attrition with high accuracy by using the most popular types of ML.

The main drawback of this approach is the black-box model, since ML models operate “as an opaque system where the internal workings of the model are not easily accessible or interpretable” (Hassija et al., 2024). While users may see the input and output, the system's decision-making process remains unclear, which may lead to bias and confusion. This lack of clarity is especially crucial when it comes down to communicating with people and making decisions about them. To deal with this problem, explainable AI (XAI) has been developed.

Explainability, or interpretability in some academic resources, in the context of AI means understanding how the model operates and aims to explain the behaviour of an otherwise opaque system. Generally, it shows what is happening in the model and provides an explanation which may be read by a human (Ali et al., 2023). There are different techniques of XAI; the main well-known ones are LIME and SHAP (SHaply Additive exPlanations).

LIME aims to explain individual predictions rather than the model as a whole. It analyses how the model behaves around a particular data point and uses a simpler model to approximate this behaviour. This allows identifying the main factors behind a prediction without requiring access to or a detailed understanding of the internal structure of the original model (Hassija et al., 2024).

SHAP is “a game-theoretic approach to explain the output of machine learning models by attributing contributions to each input feature” (Al-Ali et al., 2026). However, most SHAP-based approaches concentrate on a single-level feature and do not integrate a global analysis. To solve

this problem, Jia (2026) in her scientific paper introduced a four-level approach, proving that SHAP models may be enlarged. Models such as SHAP are especially important for HRD because they do not just show the score of the probability of leaving, but identify the core factors, such as low salary. It allows managers to take proactive actions toward battling attrition levels in real time.

The last thing that may be advanced by integrating AI chatbots is exit questionnaires. While we could not find any academic literature regarding the topic, we may suggest that, despite exit interviews, AI bots may accumulate the reasons for leaving a company across employees. A semantic approach described above may be applied to find patterns or build a word cloud with the most common reasons.

1.6. Gamedev industry

Unlike traditional IT companies, the game development industry has its own specificities, one of the most important of which is its project-based character and reliance on external funding.

In comparison with mass production, project-based development is short-term. Generally, the cycle of game production consists of three phases: pre-production (the development of a game), release, and post-production (support of a released game) (Chandler, 2009). Cycles of hiring and firing depend on the stage of the development process: as a tendency, teams are assembled in full for the active phase of game development and are then either disbanded completely or reduced to a very small core for support. This results in “flexible employment relationships and nomadic careers” (Weststar & Legault, 2024, 27).

Workers are often hired on a project basis for a 1-3-year-long period on a temporary contract or freelance basis. Generally, they have no guarantee that after completing the project they will be reassigned to the next one (Peticca-Harris et al., 2015). Thus, instead of a traditional corporate ladder with a steady climb up the grades, employees shift from one client to another in a non-linear manner, hoping to take part in some prestigious projects (Weststar & Legault, 2024, 27).

What is interesting, such instability may be observed even in companies with long-lasting successful projects: the sphere itself pushes workers to frequently change their job place to advance new skills or to get a higher salary. To illustrate, the research conducted by the International Game Development Association in 2023 (unfortunately, no research has been carried out over the last couple of years) depicts that even though most respondents claimed to be permanent workers, in the last 5 years, they had 2.2 different job places (International Game Developers Association, 2024). This poses some challenges for managers in dealing with high attrition rates.

Another unique feature of the gamedev industry is that studios (indie and even some AAA) depend on investors or publishers to produce video or mobile games, which imposes the constraints of the ‘iron triangle’ on them. The “iron triangle” means the relationships between studio managers and publishers in terms of how the budget, the deadlines, and the scope of tasks from the latter restrict the former (Cote & Harris, 2023; Weststar & Legault, 2024). It also means that teams should satisfy publishers by all means necessary. Not only that, project-based production consists of uncertainty, since in a fast-paced environment, there are no universal protocols, and they are developed and adapted as the work progresses (Weststar & Legault, 2024). Overall, it leads to a culture of crunches.

In the gamedev industry, crunches refer to working overtime to meet deadlines (Bourscheid et al., 2025). According to Cote & Harris (2023), there are two types of crunches: ‘bad’ ones, which are involuntary and not managed and lead to burnout, physical and mental health challenges and reduced time with families, and ‘good’ ones, which are managed and based on workers’ initiative. Self-imposed crunches are rooted in the passion of employees towards games, which is widespread in the industry. However, it leads to the normalisation of crunches, excessive valuation of overwork (Bourscheid et al., 2025), and the same health problems and burnout. The research by Mendes and Queiros (2020) shows that 85% of respondents have experienced crunches at least

in their career. The research of Peticca-Harris et al. (2015), conducted by analysing blogs, shows that such topics as "stress, burnout, depression, frustration, irritability, exhaustion, alcoholism, and suicidal tendencies" (Peticca-Harris et al., 2015, 20) were largely discussed by blog users under the theme of game developers' burnout. Unlikely, crunches remain endemic in the gamedev industry and require HR managers to account for the associated negative consequences.

1.7. Research gap

Literature review demonstrates a pattern that declines as the HR life cycle progresses: while the screening and recruiting stages are well covered by academic and industrial research, after that, their number decreases significantly. To be more specific, there is a significant gap in the academic literature on the implementation of AI for onboarding, engagement, and retention. It is also worth noting that most research was conducted in large companies studying ML systems that had been trained on large datasets. Furthermore, even though we were able to find a solid amount of articles studying the culture of crunch in the gamedev industry, there was no research that would connect AI in HR and game development studios.

Based on a review of the literature, this research paper proposes a model of how the role of AI bots might expand throughout an employee's entire career within small game development studios. In our case, the real numbers were allowed only at the stage of screening; thus, the model is empirically tested on this phase of the HR process. The screening stage serves as a starting point that will enable us to assess in practice the extent to which the patterns proposed in the literature (for example, reducing time-to-hire) apply to small companies and within the context of the gamedev industry. On this basis, we will be able to identify areas for further empirical research into other stages.

2. Theoretical framework

Our research paper is built through the lens of the Employee lifecycle (ELC) model. Beer et al. (1984) were the first to describe this model in their work "Managing human assets". They discovered a sequence of 8 HR processes, which may be referred to as the modern ELC: "recruitment; assessment and selection; orientation and socialisation; evaluation of performance and potential; career development; internal placement, promotion, and demotion; education and training; termination, outplacement, and retirement" (Gladka et al., 2022, 40). Gladka et al. (2022) conducted the systematic literature review and came to the conclusion that there are multiple definitions of the ELC, and the authors divide it into variously named and differentiated stages. After careful examination of the models reviewed in the article, we decided to synthesise it as follows: 'recruitment, onboarding, engagement, retention, offboarding'. As can be seen, this model has implicitly defined the scope of our literature review and can now be articulated in concrete terms.

The next step is to apply AI at every stage of the ELC. While in our literature review, we were able to demonstrate how the authors observe the introduction of AI tools in different stages of the HR process, we lacked the literature that would combine all the steps. This gap now is closed by Nosratabadi et al. (2022) who explains how AI tools may be used in each part of ELC: during recruitment phase, AI may help HR managers to match candidates' skills with requirements of a vacancy; after that, during onboarding, AI may be used to create personalised training programs; talking about engagement and retention, AI may predict employee promotion and performance and help to manage the quality of work; finally, AI may calculate the reasons for possible turnover and attrition. It is interesting to note that, whilst the article by Nosratabadi et al. (2022) confirms the potential for applying AI at all stages of the ELC, our literature review (1.1–1.6) revealed that empirical research varies across the different stages. As a framework, this model works, but the level of detail varies from phase to phase, having its maximum during the recruitment stage and levelling off to the end of the cycle.

The inconsistent application of AI is explained through another theoretical framework, which is the Technology Acceptance Model (TAM), which was introduced by Davis (1989). It states that a combination of perceived ease of use and perceived usefulness bases attitude toward usage, which, as follows, leads to behavioural intention to use and actual usage. It shows why HR managers and specialists are prepared to apply AI at different stages to varying degrees: while AI's benefits may be numerically estimated during the recruitment phase of ELC, their implications for onboarding or retention remain rather vague. Venkatesh et al. (2003) develop this model into UTAUT (Unified Theory of Acceptance and Use of Technology) and find that such factors as performance expectancy, effort expectancy, social influence, and facilitating conditions affect the readiness and the actual usage of any given technology. We expect that in small gamedev studios, the last two factors are the crucial ones when it comes to integrating AI in ELC: employees resist the implementation of AI in such delicate topics as the assessment of their engagement or retention (which correlates with the barriers proposed by Tambe et al. (2019)), and conditions of workplaces (a project-based system, culture of crunches, the dependency on investors) do not allow to actively promote new technologies throughout ELC. Kim et al. (2024) prove that point and illustrate that social influence (the support from managers and colleagues) influences the process of integration more than performance expectancy.

However, we need to underline that these models are applicable more for the voluntary usage of a technology by an individual. They are well received at the recruitment stage, where recruiters themselves are often keen to implement AI tools to speed up and simplify the process, but there is resistance at the other stages of the cycle, when HR managers wish to introduce AI tools for use by their staff. This voluntary and involuntary usage explains why the recruitment stage became the point of attraction for both academic research and practitioner implementation.

To summarise, based on the research introduced, we would like to formulate our framework as 'Differential AI Adoption Across the Employee Lifecycle'. Since the recruitment phase of ELC has the maximum voluntary adoption and the broadest scope of academic literature and industrial practices, we decided to select it as a best-case scenario and concentrate the empirical focus of our case study on it. Thus, the framework is not tested empirically in its fullest, but used as an interpretative framework for recruitment data and as a basis for formulating directions for further research into the remaining stages. Having established the framework, the following section explains how it will be empirically tested on the case of a game studio.

3. Methodology

We decided to develop our empirical research through the method of case study. The case study may be described as "an intensive, systematic investigation of a single individual, group, community or some other unit in which the researcher examines in-depth data relating to several variables" (Heale & Twycross, 2018). According to Yin (2018), it is the best option when the research question is *how* or *why* rather than *what*, focused on the process instead of measurable elements. It is also applicable when the research analyses contemporary phenomena in a real-life context, combining several data sources. Since we decided to test our 'Differential AI Adoption Across Employee Lifecycle' framework on the recruitment phase as the 'best-case scenario', it is logical to conduct a deep case study of one organisation instead of trying to make a quantitative analysis of several companies. Some research observed in the literature review (such as Sharma's study of L'Oreal (Black & Van Esch, 2020)) proves that this method is optimal for studying HR processes. To specify, in our research, we will use a mixed case study approach, since we have quantitative metrics (time-to-hire, retention, etc.) and qualitative context.

Our case study will be conducted on an anonymised game studio, to which we will refer as the 'Studio'. It is located in Eastern Europe, concentrates on producing mobile games, and consists of approximately 250-350 employees. The recruitment team is represented by fewer than 5 team members. It is necessary to mention that the author is a current employee of the Studio; hence,

they act as both a theoretician and a practitioner and, therefore, conduct an insider case study. Prior to conducting this research, the researcher has obtained the written informal permission to use anonymised, aggregate recruitment statistics from the Studio's HR Business Partner on the grounds that no personally identifiable information about candidates or employees is disclosed, and the Studio's name and location have been anonymised throughout this paper.

The main source of data is the Studio's internal quarterly HR reporting on recruitment, which was calculated manually. We will compare two time periods: before the implementation of AI, four quarters of 2025, and after the implementation, the first two quarters of 2026. Since the number of quarters is not equal, we will be comparing the quarterly averages, not their sum. There will be two parallel, non-intersecting cross-sections: aggregated figures for the entire recruitment team across all vacancies and figures for one recruiter ("Recruiter A", anonymised), specialising in creative roles and game designer positions, for the same period. The second segment was chosen to test the hypothesis about the heterogeneity of gamedev teams presented in the Introduction, since general AI learnt on tech-datasets may act differently. It is also worth noting that, in the second segment, we will have a more limited selection with fewer vacancies.

As for metrics, they may be broadly divided into three categories: process (time spent on assessing CVs, writing summaries and rejections, drawing up a candidate's profile, sourcing, preparation for screening, writing job vacancies), recruitment funnel (number of job offers, offer acceptance rate, probationary period completion rate), and retention (dismissals during the probationary period and after 3–12 months). Time-based process metrics (e.g., time spent per CV) are based on the author's retrospective estimation rather than precise time-tracking logs, since no dedicated tool was used to record this data at the time. We have the retention data in 2025, which cannot be compared with the period after implementation of AI, since not enough time has passed. As regards units of measurement, hours/minutes will be used for process metrics and percentages and absolute numbers for recruitment funnel and retention.

Given the relatively small sample size without inferential statistics, we decided to choose descriptive comparative analysis as the main method of research: quarterly averages for process metrics and a direct comparison without averaging for recruitment funnel and retention statistics. We will have two independent comparisons: team-level to test the hypothesis about recruitment as the 'best-case scenario' and role-specific to test the heterogeneity, whether this pattern will work on specific roles. If the trends in both samples are consistent, this will confirm the broad applicability of AI tools, regardless of the type of role. In contrast, if there are discrepancies, this will limit the generalisability of the model and support the claim that it is more difficult to implement AI in heterogeneous teams. Retention figures will serve as a descriptive baseline that, in the Discussion section, will form the basis for a specific hypothesis to be tested in the future. This analysis is expected to not only provide an answer on whether our framework is confirmed, but also to offer meaningful insights into its variations across different work positions.

4. Empirical Findings

4.1. AI Tools implemented in the Studio

AI tools were implemented in the Studio in January, 2026. The Studio intentionally decided not to introduce ATS systems for reviewing and scoring resumes of applicants because of its 'black-box' nature and potential bias, which correlates with the limitations we described in the literature review (1.2). Managers were sceptical about the correct assessment by these tools and proposed that they could reject too many good applicants for biased reasons or just because candidates did not put keywords into their resume. To add, while it could potentially save time for recruiters working with tech positions (Engine Developers, Go Developers, Quality Assurance Specialists, DevOps Engineers, etc.), the rest of the positions are more artistic ones. Unlike technicians, candidates for such roles as Game Designers or Project Managers, as a tendency, do not have keywords (a technological stack) in their CVs that could be easily identified by ATS. Instead, they

write about their projects, duties, and results in their own style, which demands careful observation of each resume by a recruiter. Moreover, Artists (Concept, Character, 3D, Promo, Environment, Sound Designers, etc.), Animators, Motion Designers, Sound Designers, etc. have a portfolio of their works to demonstrate their skills - AI is not able to assess the quality of a picture, a showreel, or sound, which makes it useless while working with these types of vacancies. Since the ratio between tech and more creative positions in the Studio is approximately 2:5, it was decided not to invest money into this type of AI tools and find something more applicable.

As a result, two AI assistants were developed: 'Creative Assistant' and 'Follow-up Assistant' (draft names). The AI recruiter assistant is a neural network model developed and trained using ChatGPT. It enables the automation of routine daily tasks: processing text, writing follow-ups and helping to draft emails. The difference from the basic ChatGPT is that the model already understands the general nature of the request that will be made, so there's no need to explain the task to it every time. Due to the NDA policy, we cannot provide the exact prompts, but we will describe how the agents work in detail.

Before we start describing our models, it is important to outline that any personal information should not be provided to AI, since it remains stored on external sources. It includes:

- The candidate's first name and surname;
- Just the first name, combined with any other information described below;
- Telephone numbers;
- E-mail address;
- Address (even the town, when combined with the candidate's name and places of work discussed during the interview, constitutes personal data);
- Length of service with the company;
- Any information that could identify an individual.

Any sensitive information about the Studio should not be inputted likewise. It may include:

- Any information about the unannounced projects;
- Any confidential information;
- Employees' names or surnames;
- Any financial information (budget, salaries, etc.);
- Any product metrics that are not presented on public sources.

Turning to the assistants, 'Creative Assistant' is a more creative model capable of thinking, imagining, and providing the evaluation. It has four regimes (or functions):

1. 'Cold letters' - it gathers the information about the vacancy (role, duties, project, sales points) and the candidate (the key information about their work experience and projects) and writes a personalised 'cold letter' for an initial contact with a potential candidate. The letter may be adapted for specific platforms: messaging apps, email, LinkedIn.
2. 'Questions' - based on the experience of the candidate (it is important not just to upload the CV with all personal information, but to copy-paste their work experience exclusively), description of the vacancy, requirements from the hiring managers or a list of skills that should be checked, the agent formulates the questions for a screening interview. If a recruiter wants to go deeper, they may ask the assistant to create case study questions for the specific role.
3. 'Rejection' - a recruiter writes the reasons for the rejection of a candidate, and the assistant creates a personalised rejection letter. It may be more formal or informal and include some additional details such as an offer to stay in touch. Moreover, it does not have the traditional ChatGPT's politeness and excessive praise.
4. 'Soft skills' - based on the transcript written by a recruiter during a screening interview, the assistant assesses a candidate's soft skills, behaviour and responses: where they may have fallen short and why, how they could have responded better, and what the potential risks

might be. In particular, it focuses on responsibility, independence, initiative, conflict management, reflection, flexibility, communication, attitude to feedback, discipline and people management patterns. It is important to keep in mind that AI models cannot criticise severely; thus, they formulate their opinion in a gentle manner.

'Follow-up Assistant' is a model that is, in effect, 'prohibited' from thinking and making judgements. It is a more rigorous and formal tool that does not distort the data. It also has four regimes:

1. 'Screening' - a recruiter uploads a written transcript of an interview, and an assistant summarises it following the given structure. It may also list the answers to case study questions separately. The structure may vary from recruiter to recruiter, and from position to position.
2. 'Candidate profile' - at the beginning of working on any vacancy, a recruiter has a meeting with hiring managers to specify the requirements, the tasks, and the core criteria, following which they will assess the candidates. The notes from this meeting are uploaded into a model, and it gives a structured ideal candidate's profile divided into hard skills and soft skills.
3. 'Tech Interview' - likewise the 'Screening', AI agent summarises the transcript of a technical interview conducted by hiring managers following the given structure. The model also helps to determine the final decision and the risks, which a recruiter then discusses with the team.
4. 'Boolean Search' - based on keywords in a vacancy description, an assistant creates a Boolean search query for Google. It works with such platforms as LinkedIn, ArtStation, Behance and some more, but many job platforms block their candidate bases and make them inaccessible via external sources like Google. Thus, this tool has some limitations for use.

Following the model provided by Baranyi (2025), these AI agents refer more to Digitalised Practices rather than Digitally Transformed ones, since they act as additional instruments for recruiters and do not operate automatically without human assistance.

We decided to compare the AI agent functions we used with those that had already been studied in the literature and to identify patterns. To do that, we provide the following table:

As a result, the majority of functions are described in the literature, proving that the Studio has already integrated theoretically formulated patterns of using AI in the recruitment process. The exceptions are the candidate's profile and summaries after interviews - these tools work in practice, but have not yet been explored in the academic literature.

As can be seen from the tables, there was no change in the recruitment funnel's metrics before and after the implementation of AI assistants - neither at the team level nor at Recruiter A level. There may be several factors contributing to the pattern. Firstly, we operate with two different approaches to efficiency: the speed of the process, on which AI may have an impact, and the number of closed positions, dictated by the budget, the company's annual and quarterly recruitment plans, and business requirements. The latter cannot be changed by AI. Secondly, 2026 is still continuing; thus, it is complicated to make any generalisations now. We may expect that if a recruiter has more time on their hands – thanks to the optimisation of the workflow by AI agents – they could potentially handle more vacancies at the same time. Unfortunately, at the time of writing the paper, we do not yet have data covering the entire year period, so we are leaving these comparisons for a future study.

Following the statistics described in the tables, we may see that integration of AI assistants significantly reduced the time needed for most operational work in the recruitment workflow: writing summaries, assessing resumes, writing rejections, preparing for screening interviews,

drawing up a candidate's profile, summarising the text script after screening and technical interview, and preparing Boolean queries. In all cases, time was reduced by more than 60%.
Moreover, Recruiter A, who works on more creative roles, saves more time in some instances than the whole team does (for example, -70,1% time spent on CV assessment in comparison with -66,9% for the team or -74,1% vs -72% on writing summaries), leading to the fact that custom-built AI agents (not general ATS systems built on the tech roles datasets, which the Studio refused to put in use) optimise the process despite the heterogeneity of teams, successfully dealing with creative vacancies as well.
As for retention metrics, it is not a comparison before and after AI implementation, since the number of dismissed employees was calculated throughout a 3-year-long period. The table provides a descriptive baseline for future analysis, when the data will be sufficient.
One possible future outcome may be that if AI-assisted screening improves candidate-role fit, the share of early departures (<3 months) among 2026 hires might fall below the 17.5% 2025 baseline. However, this hypothesis can only be tested once sufficient time has passed. Not only that, the last table correlates with the research by the International
Game Developers Association (2024), which showed that, typically, employees had 2.2 jobs during the last 5 years. According to the table, most workers either leave or are fired during the first year or after 1-3 years of employment.

5.3. Practical implications

Practical implications of our paper follow from the voluntary/involuntary distinction of the UTAUT (Venkatesh et al., 2003) from the theoretical framework and the typology of Baranyi (2025).
Following the logic of UTAUT, one of the key factors affecting the successful implementation of a new technology is the voluntary use of it by employees. Unlike other stages of ELC (onboarding, retention, offboarding, etc.), during which workers may show some resistance towards AI tools, in the recruitment stage of ELC, recruiters are generally more welcoming to integrating AI assistants into their workflow, because it optimises the processes and saves their time. Our research supports that statement with the considerable changes in process metrics, showing that the execution times for all functions that had been complemented with AI agents have been reduced by more than 60%. It is important to note that the Studio designed assistive instruments: instead of giving autonomy to AI agents (which practice is especially relevant with ATS systems), recruiters work with them as with supportive tools and decide for themselves in which circumstances and to what extent to use them. Moreover, there are even some more delicate functions such as the assessment of soft skills (an element of human estimation, not just an administrative function like summarising transcripts), with which AI also copes well. It shows that the voluntary framing of the instrument is more important than the actual content of the problem. What is important, this design of instruments may offer a path toward implementing AI tools at other stages of ELC where resistance is typically higher.
Talking about Baranyi's typology (2025), AI agents of the Studio preserve the 'human-in-the-loop' practice, which makes them Digitalised Practices. Whilst large studios are striving to automate every possible function, smaller companies face a dilemma: to what extent should they outsource ATS decisions to optimise time whilst avoiding staff dissatisfaction? Should they implement Digitalised or Digitally Transformed Practices? Our research answers this question: the Digitalised approach does not compromise on efficiency (reducing time by 60–77% across all process metrics) while offering two additional benefits: a voluntary usage pattern and a reduction in certain ethical risks.
Talking about ethical risks, the critique of bias and the 'black-box' nature of AI (Köchling & Wehner, 2020) is mainly addressed to the systems that operate autonomously and affect the decision-making process without clarity (like ATS systems). In our design, the final decision on each candidate rests with a recruiter. To put it as an example, AI assessment of soft skills of a candidate

is formulated in a rather 'soft' form, allowing a recruiter to make the final verdict on their own. Maintaining a 'human-in-the-loop' approach is not merely an ethical precaution 'just in case', but a tested strategy which, in our case, combines high efficiency with a low level of algorithmic risk.
To summarise, we advise small studios to make a choice for custom-built, assistive tools designed with the specific needs of particular roles and teams in mind. As the practice shows, voluntary usage and well-thought-out design show successful AI applicability even in heterogeneous teams with technical and creative roles.

Conclusion

To conclude, we would like to remind about our research questions: How does AI adoption vary across the stages of the employee lifecycle? Focusing empirically on recruitment as a best-case scenario, to what extent does this pattern hold across heterogeneous roles within a game development studio? While the first question was answered through the literature review and proposition of the 'Differential AI Adoption Across the Employee Lifecycle' framework based on it, the second one was addressed through a comparative descriptive analysis of metrics within the Studio.
Our 'Differential AI Adoption Across the Employee Lifecycle' synthesised the ELC model with the UTAUT, explaining the uneven implementation of AI tools via a voluntary/involuntary usage pattern. The empirics show that the design of integrated AI instruments (custom-built or generic) is a valuable variable, which should be taken into account in future versions of the framework.

As for the empirical contribution, a case study on the example of the Studio illustrates the considerable reduction of time across all process metrics from 60 to 77% during the recruitment stage of the ELC. Moreover, the creative roles for which Recruiter A was responsible demonstrated comparable or even better results, which refutes the claim that AI tools do not work well with this segment. However, it is important to note that in our case, custom-built AI tools were analysed - probably, the application of generic AI tools would show lower results.

Based on our research, we advise small heterogeneous gamedev studios to implement custom assistive AI solutions instead of ready-made, off-the-shelf, generalised ones to maintain effectiveness and minimise ethical risks and misunderstandings regarding the black-box model.

During our research, we have come across several limitations that should be mentioned. We conducted our research on a single-case study, analysing only one studio, which prevents our conclusions from being generalised to a certain extent. Another potential bias of interpretation is the insider role of the researcher, who is a current employee of the Studio. The period of observation after the implementation of AI assistants was short; hence, it was not possible to draw definitive conclusions based on recruitment funnel and retention metrics at the moment of writing the paper. Last but not least, our framework was tested only on the recruitment stage, which makes other phases of the ELC rather theoretical than empirical.

At the end, based on our paper, we would like to propose some areas for further research. Firstly, we are going to test our hypothesis about the effect of AI tools during recruitment on the percentage of early departures (probably, thanks to AI assistants, more suitable candidates will be selected, which will reduce the number of people dismissed during their probationary period). Secondly, we would like to replicate our case on other gamedev studios to test the generalisability of our framework and empirical findings. Thirdly, the empirical test of the framework on other stages of the ELC should be done; nevertheless, there are some limitations, since the implementation of AI tools in further steps of the ELC is beyond the researcher's remit. Finally, we would like to compare custom-built and generic AI tools directly - inside one company or across

several studios - to test the hypothesis that the design of the instrument is more important than the content of the position.

References

1. Al-Ali, M., Alwateer, M., Alsaedi, S. A., Balaha, H. M., Badawy, M., & Elhosseini, M. A. (2026). Integrating machine learning and explainable AI for employee attrition prediction in HR analytics. *Scientific Reports, 16*(1), 6344.
2. Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion, 99*, Article 101805. https://doi.org/10.1016/j.inffus.2023.101805
3. Bain & Company. (2023, October 4). *Beyond the love of the game: Talent in the video game industry*. https://www.bain.com/insights/beyond-the-love-of-the-game-talent-in-the-video-game-industry/
4. Baranyi, V. (2025). Systematic literature review on the digital transformation of the personnel selection process. *Journal of Work-Applied Management*. https://doi.org/10.1177/23970022251363012
5. Beer, M., Spector, B., & Lawrence, P. R. (1984). *Managing human assets*. Free Press.
6. Bindra, S., Bhattacharya, S., & Bhattacharya, S. (2025). Traditional to digital: Human resource management transformation. *Journal of Work-Applied Management*.
7. Black, J. S., & Van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons, 63*(2), 215–226.
8. Blount, J. B. (2022). Betting on talent: Examining the relationship between employee retention and onboarding programs. *Engaged Management Review, 5*(3). https://commons.case.edu/emr/vol5/iss3/3/
9. Bourscheid, D. P., Steil, A. V., & Paixão, E. D. S. (2025, April). Crunch on video game production: Practices to avoid for healthier workplaces in game development. In *Proceedings of the 20th International Conference on the Foundations of Digital Games* (pp. 1–5).
10. Ch, R. K., & Singha, S. (2024). A strategy for conducting sentiment analysis using AI in human resource management. In R. Khamis & A. Buallay (Eds.), *AI in business: Opportunities and limitations* (Vol. 515, pp. 491–501). Springer. https://doi.org/10.1007/978-3-031-48479-7_42
11. Chandad, A., Benchekroun, M. A., & Abakouy, M. (2026). Transformations in talent acquisition: Measuring AI's effects on recruitment efficiency and bias. *F1000Research, 15*(802), 802.
12. Chandler, H. M. (2009). *The game production handbook*. Jones & Bartlett Publishers.
13. Chen, Z. (2023). Collaboration among recruiters and artificial intelligence: Removing human prejudices in employment. *Cognition, Technology & Work, 25*(1), 135–149.
14. Cote, A. C., & Harris, B. C. (2023). The cruel optimism of "good crunch": How game industry discourses perpetuate unsustainable labor practices. *New Media & Society, 25*(3), 609–627.
15. Davis, F. D. (1989). Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption, 205*(219), 5.
16. Ekuma, K. (2024). Artificial intelligence and automation in human resource development: A systematic review. *Human Resource Development Review, 23*(2). https://doi.org/10.1177/15344843231224009

17. El-Ghoul, M., Almassri, M. M., El-Habibi, M. F., Al-Qadi, M. H., Abou Eloun, A., Abu-Nasser, B. S., & Abu-Naser, S. S. (2024). AI in HRM: Revolutionizing recruitment, performance management, and employee engagement. *International Journal of Human Resource Technology, 18*(2), 112–134.
18. El Ouakili, O. (2025). The impact of artificial intelligence (AI) on recruitment process. *Open Journal of Business and Management, 13*(2), 749–762.
19. European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union, L 2024/1689*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj
20. GamesIndustry.biz. (2025, December 19). *Games industry layoff figures were down slightly in 2025 – but it was still horrendous | Year in review*. https://www.gamesindustry.biz/games-industry-layoff-figures-were-down-slightly-in-2025-but-it-was-still-horrendous-year-in-review
21. Gelbard, R., Ramon-Gonen, R., Carmeli, A., Bittmann, R. M., & Talyansky, R. (2018). Sentiment analysis in organizational work: Towards an ontology of people analytics. *Expert Systems, 35*(5), e12289.
22. Giamos, D., Doucet, O., & Léger, P. M. (2024). Continuous performance feedback: Investigating the effects of feedback content and feedback sources on performance, motivation to improve performance and task engagement. *Journal of Organizational Behavior Management, 44*(3), 194–213. https://doi.org/10.1080/01608061.2023.2238029
23. Gladka, O., Fedorova, V., & Dohadailo, Y. (2022). Development of conceptual bases of the employee life cycle within an organization. *Business: Theory and Practice, 23*(1), 39–52.
24. Haque, A. (2025). AI is hiring you: Algorithmic power, ethical risks, and the future of recruitment. *International Journal of Organization Theory & Behavior*, 1–9. https://doi.org/10.1108/IJOTB-02-2025-0045
25. Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation, 16*, 45–74. https://doi.org/10.1007/s12559-023-10179-8
26. Heale, R., & Twycross, A. (2018). What is a case study? *Evidence Based Nursing, 21*(1), 7–8.
27. Hinge, P., Thakur, A., & Salunkhe, H. (2023, May). Analysis of human resources attrition: A thematic and sentiment analysis approach. In *International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)* (pp. 820–828). Atlantis Press. https://doi.org/10.2991/978-94-6463-136-4_72
28. IconEra. (2025, December 18). *Video game statistics 2026*. https://icon-era.com/statistics/video-game-statistics/
29. Idrus, S. (2025). Performance management reimagined: Moving beyond annual reviews to continuous feedback. *The Journal of Academic Science, 2*(1). https://doi.org/10.59613/q0e78r85
30. International Game Developers Association. (2024, March 14). *Developer satisfaction survey 2023: Summary report*. https://igda-website.s3.us-east-2.amazonaws.com/wp-content/uploads/2024/05/01161842/2023-04-14_IGSA-DSS-2023_SummaryReport.pdf
31. Ip, E. (2025). Fair AI in hiring: Experimental evidence on how biased hiring algorithms and different debiasing methods affect the quality and diversity of applicants. *Journal of Economics, Finance and Administrative Science, 11*(1). https://doi.org/10.1177/23794607251353585

32. Jasiak-Kaczmarek, B. (2025). Implementation of artificial intelligence for employee onboarding in light of literature review. *Zeszyty Naukowe. Organizacja i Zarządzanie/Politechnika Śląska*, (219), 281–292.
33. Jia, H. (2026). Explainable AI for employee turnover prediction: A SHAP-based intelligent analytics approach. *Scientific Reports*.
34. Kim, Y., Blazquez, V., & Oh, T. (2024). Determinants of generative AI system adoption and usage behavior in Korean companies: Applying the UTAUT model. *Behavioral Sciences, 14*(11), 1035.
35. Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research, 13*(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w
36. Krishnan, L. R. K., Praveen, K., & Poorani, S. (2024). Artificial intelligence in human resource management: Enhancing efficiency and transforming employee experience. *The Indian Journal of Industrial Relations: A Review of Economic and Social Development, 59*(4), 589–605. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4782399
37. Kumar, N. (2025). *AI recruitment statistics 2025 (Worldwide data and insights)*.
38. Kumar, A., & Singh, P. (2024). Natural language processing applications in HR: Designing effective chatbot systems for employee support. *AI and Machine Learning in Business, 8*(4), 423–445.
39. Leong, C. (2018). Technology & recruiting 101: How it works and where it's going. *Strategic HR Review, 17*(1), 50–52.
40. Maganti, S. (2025). From annual reviews to continuous feedback: Evolving performance appraisal systems in 2025. *International Journal of Research in Management, 7*(1), 1039–1049. https://doi.org/10.33545/26648792.2025.v7.i1k.528
41. McKinsey & Company. (2025, November 5). *The state of AI: Global survey*. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai
42. Mendes, J., & Queirós, C. (2020, November). It's crunch time: Burnout, job demands and job resources in game developers. In *International Conference on Videogame Sciences and Arts* (pp. 46–58). Springer International Publishing.
43. Nawaz, N., Arunachalam, H., Pathi, B. K., & Gajenderan, V. (2024). The adaption of artificial intelligence in human resource management practices. *International Journal of Information Management Data Insights, 4*(1). https://doi.org/10.1016/j.jjimei.2023.100208
44. Nandal, M., Grover, V., Sahu, D., & Dogra, M. (2024). Employee attrition: Analysis of data driven models. *EAI Endorsed Transactions on Internet of Things, 10*(1).
45. Nosratabadi, S., Zahed, R. K., Ponkratov, V. V., & Kostyrin, E. V. (2022). Artificial intelligence models and employee lifecycle management: A systematic literature review. *Organizacija, 55*(3), 181–198.
46. Peticca-Harris, A., Weststar, J., & McKenna, S. (2015). The perils of project-based work: Attempting resistance to extreme work practices in video game development. *Organization, 22*(4), 570–587.
47. Priyadarsini, P., & Sreejith, S. S. (2026). Identifying the most effective recruitment metrics using analytical hierarchy process. *Future Business Journal, 12*(1), 144.
48. Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI and recruitment. *Computer Law & Security Review, 53*, 105966. https://doi.org/10.1016/j.clsr.2024.105966
49. Ritz, E., Donisi, F., Elshan, E., & Rietsche, R. (2023). Artificial socialization? How artificial intelligence applications can shape a new era of employee onboarding practices. In

Proceedings of the 56th Annual Hawaii International Conference on System Sciences (pp. 155–164). https://doi.org/10.24251/HICSS.2023.019

50. Rivera, M., Qiu, L., Kumar, S., & Petrucci, T. (2021). Are traditional performance reviews outdated? An empirical analysis on continuous, real-time feedback in the workplace. *Information Systems Research, 32*(2). https://doi.org/10.1287/isre.2020.0979
51. Sheard, N. (2025). Algorithm-facilitated discrimination: A socio-legal study of the use by employers of artificial intelligence hiring systems. *Journal of Law and Society*. https://doi.org/10.1111/jols.12535
52. Slaykovskiy, V., Zvegintsev, M., Sakhonchyk, Y., & Ajamian, H. (2025). Evaluating AI recruitment sourcing tools by human preference. *arXiv preprint arXiv:2504.02463*.
53. Society for Human Resource Management. (2024, May 15). *Employee onboarding guide: Measuring success*. https://www.shrm.org/topics-tools/topics/onboarding/measuring-success
54. Sudhakar, M. (2022). *How conversational AI is changing the HR landscape*. HR.com. https://www.hr.com/en/magazines/talent_acquisition/february_2022_talent_acquisition_excellence/how-conversational-ai-is-changing-the-hrlandscape_kzr5b3id.html
55. Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42.
56. Traynor, S., Wellens, M. A., & Krishnamoorthy, V. (2021). Continuous performance management. In S. Traynor, M. A. Wellens, & V. Krishnamoorthy (Eds.), *SAP SuccessFactors talent: Volume 1: A complete guide to configuration, administration, and best practices: Performance and goals* (pp. 613–643). Apress. https://doi.org/10.1007/978-1-4842-6600-7_13
57. Upadhyay, A. K., & Khandelwal, K. (2018). Applying artificial intelligence: Implications for recruitment. *Strategic HR Review, 17*(5), 255–258.
58. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478.
59. Viterouli, M., Belias, D., Koustelios, A., Tsigilis, N., & Papademetriou, C. (2024). Time for change: Designing tailored training initiatives for organizational transformation. In *Organizational behavior and human resource management for complex work environments* (pp. 267–307). IGI Global.
60. Wandhe, D. P. (2026). *AI chatbots as enablers of effective employee onboarding*. SSRN 6099951.
61. Weststar, J., & Legault, M. J. (2024). *Not all fun and games: Videogame labour, project-based workplaces, and the new citizenship at work*.
62. WifiTalents. (2026, February 27). *Video game industry statistics*. https://wifitalents.com/video-game-industry-statistics/
63. World Economic Forum. (2025, March 28). *Hiring with AI doesn't have to be so inhumane*. https://www.weforum.org/stories/technological-innovation/ai-hiring-human-touch-recruitment/
64. Xiang, H., Lu, J., Kosov, M. E., Volkova, M. V., Ponkratov, V. V., Masterov, A. I., Elyakova, I. D., Popkov, S. Y., Taburov, D. Y., Lazareva, N. V., Muda, I., Vasiljeva, M. V., & Zekiy, A. O. (2023). Sustainable development of employee lifecycle management in the age of global challenges: Evidence from China, Russia, and Indonesia. *Sustainability, 15*(6). https://doi.org/10.3390/su15064987
65. Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

The Role of AI and IoT in Achieving Net-Zero Energy High-Rise Buildings in China

Yuan Xiaoying
Al-Farabi Kazakh National University, DBA Program Doctoral Student
Scientific Advisors:
Asanova Daniya
DBA, PhD, Professor
Olga Koshkina
c.e.s., assistant Professor

1. Abstract

China's building sector accounts for roughly 20 to 22 percent of the country's total carbon emissions and must peak alongside the broader economy by 2030 under China's dual-carbon commitments, yet high-rise buildings — a defining feature of Chinese urban form — present specific technical challenges for reaching net-zero energy performance that low-rise typologies do not. This article examines the role of artificial intelligence (AI) and the Internet of Things (IoT) in advancing net-zero energy performance in China's high-rise building stock. Drawing on recent academic literature on nearly-zero-energy buildings (NZEBs), AI-enabled building energy management, and China's building-sector decarbonization pathway, the study finds that AI and IoT contribute primarily through operational optimization — predictive HVAC control, occupancy-based automation, and digital-twin-based performance monitoring — rather than through the building envelope and on-site generation measures that dominate current Chinese NZEB policy and standards. The analysis concludes that AI and IoT are necessary but insufficient tools for high-rise net-zero performance, and are most effective when layered onto, rather than substituted for, the physical envelope and renewable-generation measures already prioritized in China's national NZEB standards.

Keywords: net-zero energy buildings, artificial intelligence, Internet of Things, high-rise buildings, China, dual carbon, building energy management, digital twin

2. Literature Review

China's building sector is one of the country's three largest energy-consuming and emitting sectors, accounting for approximately 22 percent of national carbon emissions in 2020, with building operations alone generating 2.16 billion tonnes of CO2 (City-level building operation and end-use carbon emissions dataset, 2024). Under China's 2020 "dual carbon" pledge — peaking national emissions by 2030 and achieving carbon neutrality by 2060 — the Ministry of Housing and Urban-Rural Development issued an Implementation Plan for Peak Carbon in Urban and Rural Construction in 2022, and modelling using China's Building Energy and Emission Model indicates the building sector's final energy demand is expected to peak around 2030 before declining, with direct fossil-fuel use in buildings eliminated by 2060 under a trajectory consistent with the dual-carbon targets (Regulatory Assistance Project, 2024). Scenario analysis further suggests that with an optimal combination of available technologies, building-sector carbon emissions could peak as early as 2030 at only 4.4 percent above 2020 levels, though a business-as-usual trajectory would miss the 2030 target and peak only around 2041 (Optimal carbon emission reduction path of the building sector, 2024).

China's nearly-zero-energy building (NZEB) programme has produced national and local standards and demonstration projects covering nearly 24 million square meters nationwide, alongside a

policy roadmap requiring new buildings to meet NZEB standards by 2025 and the sector to reach zero energy or carbon by 2050 (Overview of research and development of nearly zero energy buildings in China, 2024). This literature identifies incremental construction cost, the retrofitting of the existing building stock, and the limited range of current policy incentives as the principal remaining barriers to scaling NZEB adoption, but does not focus specifically on the high-rise typology, which faces additional constraints — limited roof area relative to total floor area for on-site photovoltaic generation, and greater HVAC load complexity across many floors and orientations — not shared by low-rise buildings.

A separate and rapidly growing literature addresses AI and IoT's role in building energy performance generally. A systematic review of AI and AI-powered digital twins for smart, green, and zero-energy buildings documents a sharp rise in relevant publications, from single digits in 2020 to well over 100 in 2024 across major databases, reflecting the field's rapid recent growth (AI and AI-powered digital twins for smart, green, and zero-energy buildings, 2025). AI-enabled HVAC systems are reported to reduce energy consumption through predictive control that anticipates occupancy and thermal load rather than reacting to it, and are identified as a leading application route toward zero-emission buildings specifically (Artificial intelligence-enabled heating, ventilation, and air conditioning systems, 2025). A gap in this literature is the limited attention to the specific interaction between AI/IoT optimization and the structural and typological constraints of super-tall and high-rise buildings in the Chinese context; this article addresses that gap.

3. Methodology

The research utilizes a qualitative comparative literature review:

- Policy and Standards Review: Analysis of China's national dual-carbon targets, NZEB standards, and the Ministry of Housing and Urban-Rural Development's building-sector carbon peaking plan.
- Technology Literature Review: Peer-reviewed research on AI-enabled building energy management, AI-powered digital twins, and IoT-based building automation, published between 2022 and 2026.
- Typological Gap Analysis: Comparative assessment of which net-zero energy measures (envelope, on-site generation, operational optimization) are typology-neutral versus high-rise-specific, drawing on the reviewed literature.

Because the literature specifically addressing AI/IoT applications in Chinese high-rise buildings (as opposed to buildings generally, or China's NZEB programme generally) remains limited, this article synthesizes across these three separate literatures rather than relying on a single integrated source base, and treats the resulting synthesis as indicative rather than exhaustive.

4. Results and Discussion

4.1 The High-Rise-Specific Net-Zero Challenge

Net-zero energy performance requires that a building's on-site renewable generation offset its annual energy consumption. For high-rise buildings, this equation is structurally more difficult than for low-rise typologies: available roof area for photovoltaic generation scales with a building's footprint, while total energy demand scales with its floor area across many stories, so the ratio of available on-site generation capacity to total demand falls sharply as building height increases. China's building-integrated photovoltaic (BIPV) literature, reviewed primarily in the context of lower-rise and mixed-use buildings, identifies façade-integrated photovoltaics as a partial response — extending generation surface from the roof to the building envelope — but notes that BIPV in China still faces cost, efficiency, and standardization challenges that limit its contribution relative to roof-mounted systems (Net-zero energy consumption building in China, 2023). This structural gap is precisely where operational optimization through AI and IoT becomes proportionally more important for high-rises than for low-rise buildings: since generation-side

measures alone cannot close the gap, demand-side reduction through smarter operation must carry a larger share of the burden.

4.2 AI for Predictive HVAC and Load Management

Heating, ventilation, and air conditioning typically represents the largest share of a high-rise building's operational energy demand, and is also the subsystem where AI-based optimization is most extensively documented. AI-enabled HVAC systems use predictive models — rather than purely reactive thermostatic control — to anticipate occupancy, weather, and thermal load, adjusting output ahead of demand rather than in response to it, an approach systematically reviewed as a leading application route toward zero-emission building performance (Artificial intelligence-enabled heating, ventilation, and air conditioning systems, 2025). Occupancy-based automation extends this logic building-wide, using sensors to detect presence and adjust lighting, heating, and cooling accordingly, while predictive-maintenance algorithms anticipate equipment failures before they occur, reducing both energy waste from malfunctioning equipment and the emissions embodied in premature equipment replacement (SolarTech Energy Systems, 2025).

4.3 IoT Sensing and AI-Powered Digital Twins

IoT sensing provides the continuous, granular data — temperature, occupancy, air quality, equipment status — that AI-based optimization requires, and its integration with digital twin technology extends this capability from single-subsystem optimization to whole-building performance simulation and monitoring. A systematic review of AI-powered digital twins for zero-energy buildings finds this combination increasingly applied to simulate building performance under varying conditions, identify inefficiencies invisible to conventional building management systems, and support ongoing operational decision-making rather than only design-stage analysis (AI and AI-powered digital twins for smart, green, and zero-energy buildings, 2025). For high-rise buildings specifically, this capability is valuable in proportion to the building's internal complexity: a fifty-story tower has substantially more zones, orientations, and occupancy patterns to reconcile simultaneously than a single-story building, making a unified digital-twin view of the whole structure disproportionately more valuable as height and floor count increase.

4.4 Barriers and the Limits of Operational Optimization

Despite this potential, the literature is explicit that operational AI/IoT optimization cannot substitute for envelope and generation-side measures; it can only improve the efficiency with which a given physical building converts input energy into occupant comfort. Workshop-based analysis convened by the American Council for an Energy-Efficient Economy concluded that AI holds "vast potential" for accelerating progress toward net-zero buildings but identified significant implementation challenges, without suggesting that AI alone is sufficient to reach net-zero performance in the absence of parallel envelope and generation investment (ACEEE, 2024). For China's high-rise stock specifically, this implies that AI and IoT investment should be sequenced as a complement to, not a substitute for, continued progress on China's NZEB envelope standards and BIPV deployment — an especially important point given that China's own NZEB literature identifies incremental construction cost as a primary barrier to broader adoption of the physical measures that operational optimization alone cannot replace (Overview of research and development of nearly zero energy buildings in China, 2024).

5. Conclusion

This article has examined the role of AI and IoT in advancing net-zero energy performance for China's high-rise building stock, situating that role within China's dual-carbon building-sector targets and its existing NZEB standards programme. The evidence indicates that AI and IoT contribute most directly through operational optimization — predictive HVAC control, occupancy-based automation, and digital-twin-based whole-building monitoring — and that this contribution is disproportionately valuable for high-rise buildings specifically, both because their on-site

generation potential is structurally constrained relative to low-rise typologies and because their internal operational complexity is correspondingly greater.

The central finding, however, is that AI and IoT are necessary but not sufficient tools for high-rise net-zero performance. Because generation-side measures such as building-integrated photovoltaics cannot alone close the gap between a tall building's available roof and façade area and its total energy demand, operational optimization must be paired with continued investment in envelope performance and on-site generation technology rather than treated as a substitute for it. For Chinese policymakers and developers, this suggests that AI/IoT investment in high-rise net-zero projects should be planned and evaluated alongside — not instead of — the physical building measures already prioritized in China's national NZEB standards, with the two categories of investment treated as complementary components of a single decarbonization strategy rather than competing options.

References

1. ACEEE. (2024). *Can artificial intelligence get us to net-zero buildings?* Retrieved from https://www.aceee.org/blog-post/2024/11/can-artificial-intelligence-get-us-net-zero-buildings
2. Frontiers in Energy Efficiency. (2026). A review of prediction models of total carbon emission for civil buildings in China. *Frontiers*. Retrieved from https://www.frontiersin.org/journals/energy-efficiency/articles/10.3389/fenef.2023.1219776/full
3. Integrating IoT and AI for sustainable energy-efficient smart building: Potential, barriers and strategic pathways. (2025). *Sustainability, 17*(22), 10313. https://doi.org/10.3390/su172210313
4. Overview of research and development of nearly zero energy buildings in China. (2024). *National Science Open.* Retrieved from https://www.nso-journal.org/articles/nso/full_html/2024/03/NSO20230083/NSO20230083.html
5. Regulatory Assistance Project. (2024). *What is a feasible pathway to full building sector decarbonisation for China?* Retrieved from https://www.raponline.org/blog/building-sector-decarbonisation-china/
6. Optimal carbon emission reduction path of the building sector: Evidence from China. (2024). *Science of the Total Environment.* Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0048969724006909
7. City-level building operation and end-use carbon emissions dataset from China for 2015–2020. (2024). *Scientific Data.* Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10817938/
8. SolarTech Energy Systems. (2025). *Green building trends 2025: Complete guide to sustainable construction.* Retrieved from https://solartechonline.com/blog/green-building-trends-2025-guide/
9. AI and AI-powered digital twins for smart, green, and zero-energy buildings: A systematic review of leading-edge solutions for advancing environmental sustainability goals. (2025). *ScienceDirect.* Retrieved from https://www.sciencedirect.com/science/article/pii/S2666498425001061
10. Net-zero energy consumption building in China: An overview of building-integrated photovoltaic case and initiative toward sustainable future development. (2023). *Buildings, 13*(8), 2024. https://doi.org/10.3390/buildings13082024
11. Artificial intelligence-enabled heating, ventilation, and air conditioning systems toward zero-emission buildings: A systematic review of applications, challenges, and future

directions. (2025). *Applied Sciences, 15*(19), 10497. https://doi.org/10.3390/app151910497

12. AI-driven innovations in building energy management systems: A review of potential applications and energy savings. (2024). *Energies, 17*(17), 4277. https://doi.org/10.3390/en17174277

DESIGNING AND PRETESTING A SERVICE SATISFACTION QUESTIONNAIRE FOR ELDERLY HOME-BASED CARE IN CHINA: AN INSTRUMENT DEVELOPMENT PROTOCOL

Qiao Liang
Al-Farabi Kazakh National University, DBA Program Doctoral Student
Scientific Advisors:
Kozin Valeriy Aleksandrovich
MBA, Senior Data Analyst
Abinova Alfiya Yurievna
PhD, Associate Professor

Abstract

Structural equation modeling of elderly-care satisfaction depends on a measurement instrument that is both psychometrically sound and usable by respondents who may have limited literacy, reduced stamina for long interviews, and unfamiliarity with abstract rating tasks. This paper sets out the design principles, structural logic, and pretesting protocol for a service-satisfaction questionnaire built around an integrated Structure-Process-Outcome and Customer-Satisfaction-Index (SERV-SPO-CSI) framework, intended for use with elderly recipients of home-based, community, and institutional care in China. The paper describes the questionnaire's four-module structure, the item-generation and simplification procedures applied to make it accessible to older respondents, and the planned pilot-testing and psychometric-refinement criteria — reliability screening by Cronbach's alpha and exploratory factor analysis — that will be used to finalize the instrument before full-scale data collection.

Keywords: questionnaire design, instrument development, elderly care, service satisfaction, pilot study, Cronbach's alpha, exploratory factor analysis, China.

Introduction

China's rapidly expanding elderly-care system — organized around the “9073” pattern in which the large majority of older adults age at home, a smaller share use community day services, and the remainder reside in institutions — requires reliable instruments for measuring how satisfied care recipients are with the services they receive. Structural equation modeling of the kind required to test an integrated service-quality model places heavy demands on the underlying questionnaire: items must map cleanly onto latent constructs, demonstrate acceptable reliability and validity, and still be answerable by respondents who are often in their seventies or eighties, some with limited formal education, reduced hearing or eyesight, or fatigue during longer interviews. A questionnaire built primarily for a working-age or ordinarily literate population cannot simply be reused for this group without a dedicated adaptation and pretesting process. This paper presents the design principles and a stepwise pretesting protocol for such an instrument, developed for use in an empirical study of elderly-care satisfaction in Shanxi Province.

Purpose and Design Principles of the Instrument

The instrument is designed to operationalize the SERV-SPO-CSI conceptual model, which links structural conditions of care provision, perceived process quality along SERVQUAL-type

dimensions, and overall satisfaction and behavioral outcomes. Three design principles guide item construction. First, construct correspondence: every item must be traceable to a specific construct in the model (a structural antecedent, a process-quality dimension, or a satisfaction/outcome indicator) so that the resulting data support confirmatory measurement and structural testing rather than an undifferentiated satisfaction score. Second, respondent accessibility: items are written in short, concrete sentences, avoiding double-barreled questions, hypothetical framing, and abstract terminology, in recognition of the cognitive and literacy profile of the target population. Third, cultural and institutional fit: item content reflects the specific service modalities present in Chinese home-based and community elderly care — assistance with meals, bathing, cleaning, and medical referral — rather than generic hospitality-sector service items borrowed unmodified from Western retail SERVQUAL applications.

Questionnaire Structure

The instrument is organized into four modules. The first module collects basic demographic and service-context information: age, gender, urban or rural residence, service type received (home-based, community day-center, or residential institution), and self-reported health or functional-dependency status. The second module measures structural antecedents as perceived by the respondent or, where more appropriate, reported by the responsible institution — staffing levels, facility conditions, and the availability of medical-care integration. The third module operationalizes perceived process quality across the five SERVQUAL-type dimensions of tangibles, reliability, responsiveness, assurance, and empathy, adapted to describe concrete caregiving behaviors rather than abstract service attributes. The fourth module measures overall satisfaction and downstream behavioral intentions, including willingness to continue using the service and willingness to recommend it to others. All substantive items use a five-point Likert scale (1 = completely disagree, 5 = completely agree), chosen over longer scales to reduce the cognitive burden of fine-grained discrimination on older respondents.

Content and Face Validity Procedures

Before pilot administration, the item pool undergoes two rounds of qualitative review. An expert panel, comprising researchers in service-quality measurement and practitioners from elderly-care institutions, reviews each item for construct relevance and wording clarity, and flags items that are redundant, ambiguous, or culturally inappropriate. A separate cognitive pretesting round is then conducted with a small number of elderly volunteers who are asked to read each item aloud and explain in their own words what they believe it is asking, which surfaces wording that is technically correct but not intuitively understood by the target population. Items are revised, simplified, or removed based on this feedback, and the sequence of items is adjusted so that simpler, more concrete questions appear first and more evaluative or comparative questions appear later, in order to build respondent confidence before the survey becomes cognitively demanding.

Pilot Testing Protocol

Following content and face validation, the revised instrument is administered to a pilot sample of 50 to 100 elderly respondents, deliberately drawn to cover the range of service types (home-based, community, institutional) and age bands present in the target population, since item performance can differ across these subgroups. Administration is conducted face-to-face rather than through self-completion, given the literacy and vision constraints common among elderly respondents; a trained interviewer reads each item aloud, records the respondent's chosen scale point, and notes any items that repeatedly prompt confusion, requests for repetition, or refusal to answer. Informed consent is obtained from each participant prior to the interview, with an explicit statement that responses will be used only for academic research and that participation is voluntary and can be withdrawn at any point.

Planned Psychometric Refinement Criteria

Once pilot data are collected, the instrument is refined using two complementary criteria. Reliability is assessed by computing Cronbach's alpha for each hypothesized subscale; items whose removal would raise the subscale's alpha, or whose item-total correlation falls below conventional thresholds, are flagged for review, and subscales with an alpha below 0.7 are treated as requiring revision rather than retention as-is. Validity is assessed through exploratory factor analysis, examining whether items load as expected onto their intended factors; items with a factor loading below 0.5, or that cross-load substantially onto a factor other than the one they were designed to measure, are candidates for removal or rewording. These statistical criteria are applied alongside, rather than instead of, the qualitative feedback gathered during pilot administration, since an item can be statistically well-behaved while still being poorly understood by respondents, or vice versa.

Expected Contribution

This protocol contributes a replicable, respondent-appropriate procedure for developing elderly-care satisfaction instruments in the Chinese context, an area where much existing survey work either adapts general-population service-quality scales without dedicated pretesting for older respondents, or relies on institution-reported outcome measures that do not capture the care recipient's own perception of quality. By tying every stage of instrument development — item generation, expert and cognitive review, pilot administration, and psychometric refinement — explicitly back to the SERV-SPO-CSI conceptual model, the resulting questionnaire is positioned to support not only descriptive reporting of satisfaction levels but also the structural equation modeling needed to test the causal pathways the model proposes.

Conclusion

A statistically sound structural model is only as credible as the instrument used to measure its constructs, and elderly respondents present measurement challenges that a generic satisfaction survey does not address. This paper has presented the design logic and a stepwise pretesting protocol — expert review, cognitive interviewing, pilot administration to 50–100 respondents, and psychometric refinement using Cronbach's alpha and exploratory factor analysis — for a questionnaire built to measure structural, process, and outcome constructs among elderly care recipients in China. Once finalized through this protocol, the instrument is intended to support a full-scale survey of no fewer than 600 respondents in Shanxi Province, providing the measurement foundation for subsequent structural equation modeling of the SERV-SPO-CSI framework.

References

1. Cronbach L.J. Coefficient Alpha and the Internal Structure of Tests // Psychometrika. — 1951. — Vol. 16, No. 3. — P. 297–334.
2. DeVellis R.F. Scale Development: Theory and Applications. — 4th ed. — Thousand Oaks, CA: SAGE Publications, 2017.
3. Netemeyer R.G., Bearden W.O., Sharma S. Scaling Procedures: Issues and Applications. — Thousand Oaks, CA: SAGE Publications, 2003.
4. Parasuraman A., Zeithaml V.A., Berry L.L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality // Journal of Retailing. — 1988. — Vol. 64, No. 1. — P. 12–40.
5. Donabedian A. Evaluating the Quality of Medical Care // The Milbank Memorial Fund Quarterly. — 1966. — Vol. 44, No. 3. — P. 166–206.
6. Law of the People's Republic of China on the Protection of the Rights and Interests of the Elderly (as amended). — Standing Committee of the National People's Congress, 1996–2018.

7. 14th Five-Year Plan for the Development of National Aging Programs and the Elderly Care Service System. — State Council of the People's Republic of China, 2022.
8. Opinions on Advancing the Construction of the Basic Elderly Care Service System. — General Office of the Communist Party of China Central Committee and General Office of the State Council, 2023.

Digital Financial Literacy and Investment Decision-Making among Young Adults in Kazakhstan

Kabyl Akzhunis

Abstract. The rapid digitalization of financial services has increased young adults' access to investment platforms, financial information, and digital financial products. At the same time, limited digital financial literacy may expose inexperienced investors to poor investment choices, misleading online information, and financial risks. This study examines the relationship between digital financial literacy and investment decision-making among young adults in Kazakhstan. A quantitative experimental design with pre-test and post-test assessments was employed, involving an experimental group and a control group. Digital financial literacy was assessed through financial knowledge, digital financial skills, awareness of digital financial risks, and the ability to evaluate online financial information, while investment decision-making was measured through scenario-based tasks related to risk, return, diversification, and information reliability. The results indicated that the experimental group demonstrated significant improvements in both digital financial literacy and investment decision-making after the intervention, whereas changes in the control group were limited. Correlation and regression analyses further showed that higher digital financial literacy was positively associated with better investment decision quality and significantly predicted improvements in investment behavior. The findings highlight the importance of integrating digital financial education into programs targeting young adults in Kazakhstan. Such initiatives may strengthen young investors' ability to assess financial risks, critically evaluate digital information, and make more informed and responsible investment decisions.

Keywords: digital financial literacy, investment decision-making, young adults, financial education, digital finance, Kazakhstan

Introduction. In recent years, the digital transformation in the financial sector has significantly changed the way individuals use financial services and make investment decisions. The widespread use of mobile banking, digital payment systems, financial technologies (FinTech), online brokerage platforms and digital investment vehicles has significantly simplified the ability to participate in the financial market. Whereas in the past, it was necessary to turn to specialized financial intermediaries or professional advisors to carry out investment transactions, today users can obtain information about securities, investment funds and other financial instruments through mobile applications and make investment decisions independently. While such changes have increased access to financial markets, they have also increased the need for new digital competencies, in addition to traditional financial knowledge, to effectively and safely use the digital financial environment.

In this context, digital financial literacy (DFL) is becoming one of the important factors explaining financial behavior. Digital financial literacy is not limited to the ability to use mobile banking or financial applications from a technical perspective. It is a multidimensional concept that combines financial and digital literacy, understanding digital financial products, assessing their risks, verifying the reliability of information, protecting personal data, and making informed financial decisions. Ravikumar et al. (2022) developed a multidimensional model to measure digital financial literacy, which included financial and digital knowledge, awareness of digital financial services, identifying and managing digital risks, knowledge of consumer rights, practical skills, and

decision-making abilities. The authors emphasize that access to digital financial services alone does not guarantee their effective use, and that insufficient DFL can lead to financial losses and digital risks (Ravikumar et al., 2022).

Digital financial literacy is especially important for young adults. As young people are a group that actively uses digital technologies, they are more likely to obtain financial information from social networks, mobile applications, and online platforms. Furthermore, while the ease of digital platforms and the speed of investment transactions may reduce barriers to entry into financial markets, they may also increase the likelihood of impulsive decision-making, risk misjudgment, and information bias in the face of insufficient financial knowledge. Bhat et al. (2024) found that digital financial knowledge, digital financial experience, and digital skills are associated with an individual's self-control and impulsive behavior. This finding suggests that digital financial literacy encompasses not only the ability to use technology, but also behavioral components of financial decisions.

Recent empirical research suggests that financial and digital financial literacy affect the quality of investment decisions. Kumar et al. (2023) examined data from 634 investors and found that digital financial literacy, financial capability, financial autonomy, and impulsivity were important factors in financial decision-making. The study also found that financial capability partially mediated the relationship between digital financial literacy and financial decision-making. Abdallah, Tfaily, and Harraf (2024) found a statistically significant relationship between digital financial literacy and financial behavior, specifically showing that financial knowledge, awareness, and decision-making skills significantly influence financial behavior. In their model, the correlation coefficient between DFL and financial behavior was 0.542, indicating that the model explained a significant portion of the variation in financial behavior.

Research on investment decisions also supports this trend. Shroff, Paliwal, and Dewasiri (2024) analyzed data from 384 investors in emerging markets and found that financial literacy levels positively influenced investment decisions. Investors with higher levels of financial literacy were more likely to make informed investment decisions. The authors link these findings to the need to develop financial education programs in emerging economies. This relationship is further complicated in the digital investment environment. Sang (2026) found in a study involving 505 individual investors that digital financial literacy not only shapes investment attitudes but also has a direct positive impact on investment decision-making.

In the context of Kazakhstan, interest in studying the relationship between financial literacy and investment decisions has been growing in recent years. Bayakhmetova et al. (2023) showed, based on a survey of the population of Kazakhstan, that there is a relationship between the financial literacy index and the effectiveness of investment decisions. The results of the study revealed that people with high financial literacy are more likely to have investment experience, and the effectiveness of investment decisions varies depending on socio-demographic factors such as age, education level and personal income. This work is important because it empirically demonstrates the relationship between financial literacy and investment behavior in Kazakhstan.

Nevertheless, several important research gaps remain in the existing literature. Research in Kazakhstan has mainly focused on traditional financial literacy, while digital financial literacy, which combines financial knowledge and digital competence, has not been sufficiently studied as a separate research subject. Although international studies have shown the impact of DFL on financial behavior, most of them are based on data from India, Kuwait and other emerging markets; empirical evidence for Kazakhstani youth is limited.

The purpose of this study is to experimentally assess the impact of digital financial literacy on the quality of investment decision-making among young adults in Kazakhstan. The study compares the results of an experimental group that participated in a digital financial education intervention with a control group, and assesses changes in digital financial literacy and investment

decision-making indicators before and after the intervention. This approach allows the study to examine the potential causal effect of increasing digital financial literacy on the quality of investment decisions, rather than simply identifying the relationship between DFL and investment decision-making.

Materials and Methods. This study employed a quantitative experimental design to examine the relationship between digital financial literacy and investment decision-making among young adults in Kazakhstan. The study involved young adults aged 18–29 years. Participants were divided into experimental and control groups, and data were collected using pre-test and post-test assessments.

At the initial stage, participants completed a pre-test measuring their level of digital financial literacy and investment decision-making ability. The experimental group then received a short digital financial literacy intervention covering financial risk and return, diversification, digital investment platforms, cybersecurity, online financial fraud, and the evaluation of financial information obtained from social media.

Following the intervention, both groups completed a post-test to assess changes in digital financial literacy and investment decision-making.

Digital financial literacy was assessed through four main dimensions: digital financial knowledge, digital financial skills, digital financial risk and security awareness, and online financial information evaluation. The items were measured using a five-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Investment decision-making was assessed through items related to risk evaluation, expected return, diversification, investment horizon, and the reliability of financial information. Higher scores indicated more informed and rational investment decisions.

Demographic characteristics, including age, gender, education, income, and previous investment experience, were also included in the analysis.

Descriptive statistics were used to summarize the characteristics of the participants and the main study variables. The reliability of the measurement scales was assessed using Cronbach's alpha.

Paired-samples t-tests were applied to compare pre-test and post-test results, while independent-samples t-tests were used to identify differences between the experimental and control groups. Pearson correlation analysis was conducted to examine the relationship between digital financial literacy and investment decision-making. Multiple regression analysis was used to assess the effect of digital financial literacy on investment decision-making. Statistical significance was established at $p < 0.05$.

Results. A total of 200 observations were included in the analysis, with 100 participants in the experimental group and 100 participants in the control group. Table 1 presents the descriptive statistics for digital financial literacy (DFL) and investment decision-making (IDM) before and after the intervention.

Table 1. Descriptive statistics of the main study variables

Variable	Experimental group, M (SD)	Control group, M (SD)
Pre-test DFL	3.15 (0.42)	3.16 (0.47)
Post-test DFL	3.77 (0.52)	3.23 (0.50)
Pre-test IDM	3.07 (0.44)	3.00 (0.48)
Post-test IDM	3.62 (0.54)	3.06 (0.49)

At the pre-test stage, the experimental and control groups demonstrated similar levels of digital financial literacy. The mean DFL score was 3.15 (SD = 0.42) in the experimental group and

3.16 (SD = 0.47) in the control group. The difference was not statistically significant, $t(198) = -0.16$, $p = .874$. Similarly, no significant baseline difference was observed in investment decision-making, $t(198) = 1.08$, $p = .282$. Thus, the two groups were comparable before the intervention.

The reliability of the measurement scales was satisfactory. Cronbach's alpha was 0.82 for digital financial literacy and 0.79 for investment decision-making, indicating acceptable internal consistency.

Following the intervention, the mean DFL score in the experimental group increased from 3.15 to 3.77, whereas the control group showed only a slight increase from 3.16 to 3.23. This pattern is presented in Figure 1.

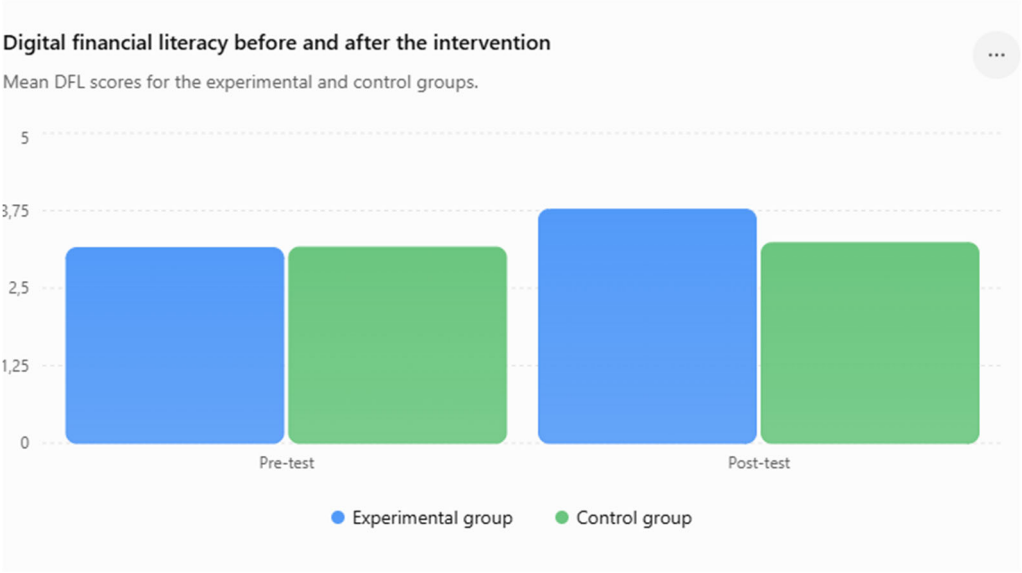
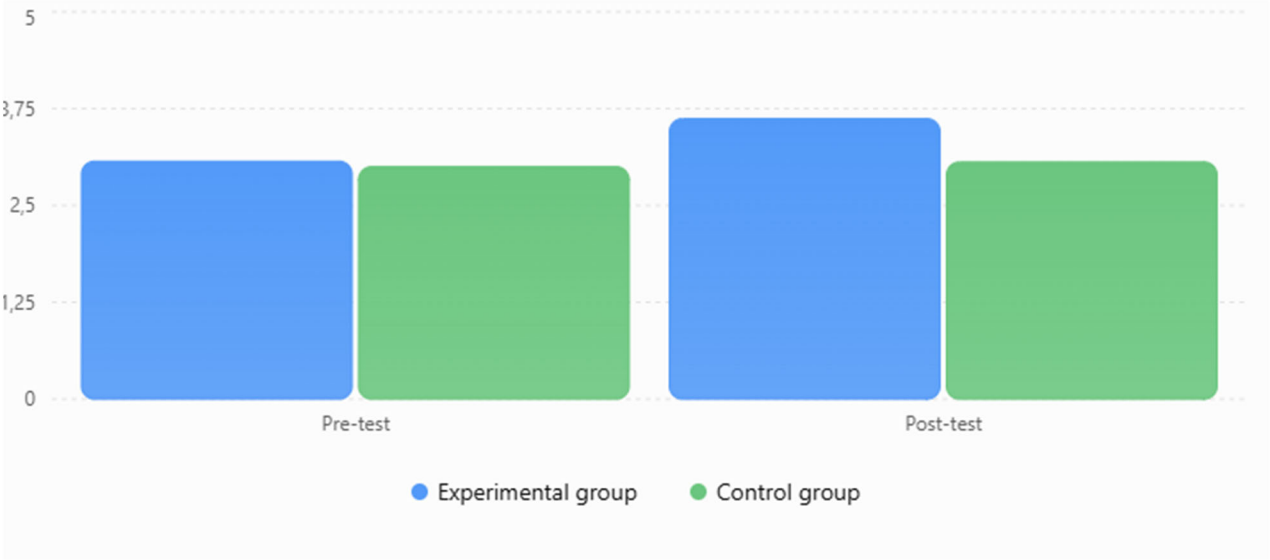


Figure 1. Digital financial literacy before and after the intervention

The results indicate a clear improvement in digital financial literacy among participants who received the intervention. In contrast, little change was observed in the control group.

A similar pattern was identified for investment decision-making. The experimental group's mean IDM score increased from 3.07 to 3.62, while the control group increased only from 3.00 to 3.06 (Figure 2)



.

Figure 2. Investment decision-making before and after the intervention

These descriptive findings suggest that the intervention was associated with improvements in both digital financial literacy and investment decision-making.

Paired-samples t-tests were conducted to examine changes between the pre-test and post-test measurements.

The experimental group demonstrated a statistically significant increase in digital financial literacy, $t(99) = 13.48$, $p < .001$, Cohen's $d = 1.35$. In contrast, the change observed in the control group was not statistically significant, $t(99) = 1.75$, $p = .083$.

Investment decision-making also improved significantly in the experimental group, $t(99) = 11.46$, $p < .001$, Cohen's $d = 1.15$, whereas the control group demonstrated no statistically significant change, $t(99) = 1.54$, $p = .127$.

Table 2. Pre-test and post-test comparisons

Outcome	Group	Mean change	t	p	Cohen's d
DFL	Experimental	+0.62	13.48	< .001	1.35
DFL	Control	+0.07	1.75	.083	0.18
IDM	Experimental	+0.55	11.46	< .001	1.15
IDM	Control	+0.06	1.54	.127	0.15

Between-group comparisons further confirmed these findings. The experimental group achieved a significantly higher post-test DFL score than the control group, $t(198) = 7.49$, $p < .001$, Cohen's $d = 1.06$. The post-test IDM score was also significantly higher in the experimental group, $t(198) = 7.68$, $p < .001$, Cohen's $d = 1.09$. Therefore, H2 and H3 were supported.

Pearson correlation analysis was performed to examine the association between digital financial literacy and investment decision-making. The analysis revealed a statistically significant positive relationship between the two variables, $r = .62$, $p < .001$.

This relationship is illustrated in Figure 3.

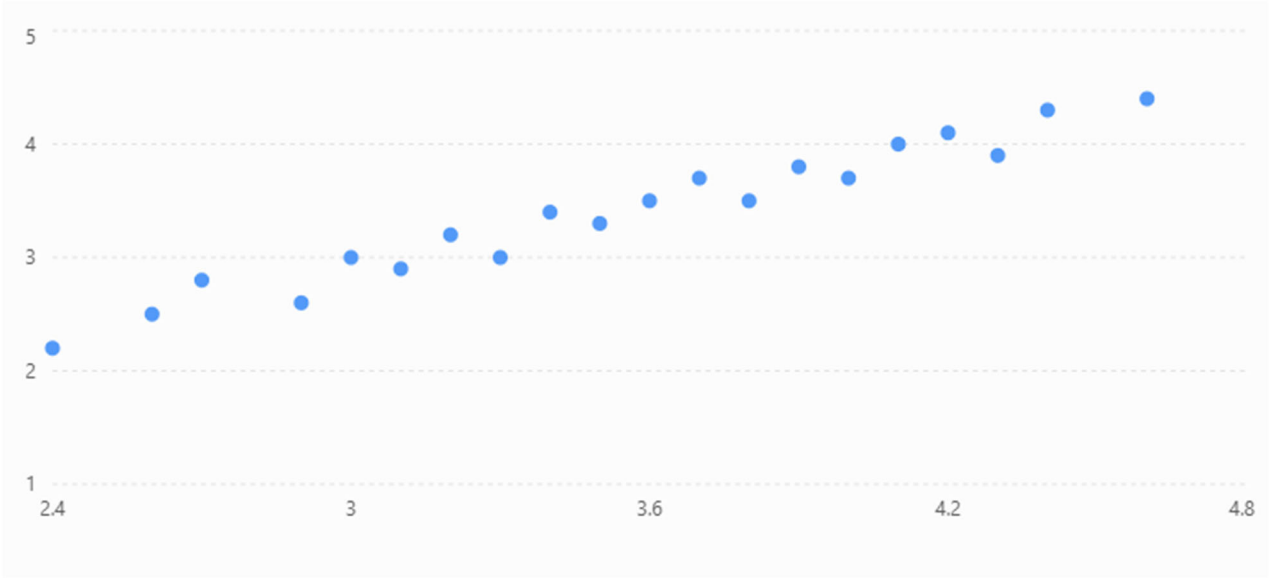


Figure 3. Relationship between digital financial literacy and investment decision-making

The upward pattern indicates that participants with higher levels of digital financial literacy generally achieved higher investment decision-making scores. The result therefore provides support for H1.

Multiple regression analysis was conducted to determine whether improvement in digital financial literacy predicted improvement in investment decision-making after controlling for demographic characteristics and previous investment experience.

The regression model was statistically significant, $F(6, 193) = 16.56$, $p < .001$, and explained 34% of the variance in changes in investment decision-making ($R^2 = .34$).

Table 3. Regression analysis predicting improvement in investment decision-making

Predictor	B	SE	β	t	p
Change in DFL	0.51	0.06	.48	8.50	< .001
Age	−0.01	0.01	−.05	−0.84	.402
Gender	0.04	0.04	.05	0.99	.324
Education	0.03	0.03	.06	1.10	.274
Income	0.05	0.03	.09	1.70	.091
Previous investment experience	0.11	0.05	.14	2.17	.031

Improvement in digital financial literacy was the strongest predictor of improvement in investment decision-making (β = .48, $p < .001$). Previous investment experience also had a small but statistically significant positive effect (β = .14, $p = .031$). Age, gender, education, and income were not statistically significant predictors.

Table 4. Summary of hypothesis testing

Hypothesis	Statistical evidence	Result
H1: DFL positively affects investment decision-making	$r = .62$, $p < .001$	Supported
H2: The intervention improves DFL	$t(99) = 13.48$, $p < .001$	Supported
H3: The intervention improves investment decision-making	$t(99) = 11.46$, $p < .001$	Supported
H4: Improvement in DFL predicts improvement in IDM	β = .48, $p < .001$	Supported

Overall, the findings indicate that the digital financial literacy intervention was associated with substantial improvements in both digital financial literacy and investment decision-making. In addition, higher digital financial literacy was positively related to better investment decision quality.

Discussion. The present study examined the effect of digital financial literacy (DFL) on investment decision-making among young adults in Kazakhstan. Overall, the findings indicate that higher levels of digital financial literacy are associated with better investment decision-making and that a targeted educational intervention can improve both DFL and investment decision quality. The significant differences observed between the experimental and control groups provide support for the proposed relationship and suggest that digital financial competencies can be developed through relatively short and focused educational interventions.

The positive association between digital financial literacy and investment decision-making supports H1. Participants with higher DFL scores tended to demonstrate more informed investment choices, particularly in relation to risk assessment, diversification, and the evaluation of financial information. This finding is consistent with Kumar et al. (2023), who identified digital financial literacy as an important determinant of financial decision-making and demonstrated that financial capability plays an important role in translating knowledge into financial behavior. Similarly, a recent meta-analysis of 30 studies found an overall positive relationship between financial literacy and investment decision-making, particularly in relation to informed investment choices and diversification.

The present findings are also consistent with Shroff et al. (2024), who reported that financially literate investors in an emerging-market context were more likely to make informed investment decisions. Their study demonstrated a positive effect of financial literacy on investment decisions among 384 investors. However, the current study extends this perspective by focusing specifically on **digital financial literacy**, which incorporates not only conventional financial knowledge but also digital skills, awareness of online financial risks, cybersecurity, and the ability to critically evaluate financial information available through digital channels.

The significant improvement in DFL scores among participants in the experimental group supports H2 and indicates that digital financial literacy is a competence that can be strengthened

through targeted education. This finding is particularly important because DFL is multidimensional rather than limited to knowledge of financial concepts. Ravikumar et al. (2022) identified financial knowledge, digital knowledge, awareness of digital financial risks, risk-control skills, consumer-rights awareness, and practical decision-making as important dimensions of digital financial literacy. Therefore, financial education designed for young adults should address not only concepts such as interest rates and diversification but also the practical challenges associated with the use of digital financial platforms.

The improvement in investment decision-making following the intervention supports H3. Participants who received digital financial literacy training demonstrated greater improvement in investment decision quality than those in the control group. This suggests that financial knowledge can be translated into more rational investment behavior when educational content is linked to practical situations involving risk, return, diversification, online information, and digital investment platforms. Rather than teaching financial concepts in isolation, scenario-based financial education may help young adults apply their knowledge when evaluating actual investment alternatives.

The regression results further support this interpretation. Improvement in DFL emerged as the strongest predictor of improvement in investment decision-making, supporting H4. The finding suggests that the relationship observed between the two variables cannot be explained solely by demographic characteristics. Previous investment experience also demonstrated a smaller positive effect, indicating that practical exposure to investment markets may contribute to decision-making competence. Nevertheless, the stronger contribution of DFL highlights the importance of knowledge and skills even among individuals with different levels of investment experience.

The findings also contribute to the growing literature on investment behavior in digital environments. Sang (2026) found that digital financial literacy directly supported informed investment decision-making but also showed that social-media influencers, impulsivity, and FinTech self-efficacy can influence investment behavior. This is particularly relevant for young investors, who increasingly obtain financial information through digital and social-media channels. Consequently, digital financial literacy should involve the ability to distinguish reliable financial information from promotional or misleading content rather than simply knowing how to access investment platforms.

The results have particular relevance to Kazakhstan. Previous research by Bayakhmetova et al. (2023) established a positive relationship between financial literacy and investment decision-making among residents of Kazakhstan and showed that investment behavior was also related to age, education, and income. The present study extends this line of research by shifting the focus from traditional financial literacy toward digital financial literacy and by concentrating specifically on young adults. More importantly, whereas the earlier Kazakhstan study relied on a questionnaire and correlational analysis, the present experimental framework evaluates changes before and after a financial literacy intervention.

From a practical perspective, the results indicate that universities, financial institutions, and policymakers in Kazakhstan may benefit from incorporating digital investment literacy into financial education programs for young adults. Such programs should combine conventional financial concepts with practical topics such as digital investment platforms, portfolio diversification, cybersecurity, financial fraud, and critical evaluation of investment advice from social media. Short scenario-based interventions may be particularly useful because they allow participants to apply financial knowledge to realistic investment decisions rather than simply memorize theoretical concepts.

Conclusion. This study examined the role of digital financial literacy in shaping investment decision-making among young adults in Kazakhstan. The findings indicate that higher levels of digital financial literacy are associated with more informed and rational investment decisions.

Participants who demonstrated stronger digital financial knowledge, better awareness of online financial risks, and greater ability to evaluate digital financial information also tended to perform better in investment-related decision-making.

The experimental results further suggest that targeted digital financial literacy education can improve both financial competence and investment decision quality. Participants exposed to the educational intervention showed greater improvements than those in the control group, particularly in areas related to risk assessment, diversification, and the evaluation of financial information. These results support the view that digital financial literacy is not limited to the technical use of financial platforms but also includes the ability to critically assess financial information and manage investment-related risks.

The study contributes to the existing literature by extending research on financial literacy in Kazakhstan toward a broader digital financial literacy perspective and by focusing specifically on young adults. The findings also highlight the potential value of practical financial education programs that combine traditional financial knowledge with digital skills, cybersecurity awareness, and critical evaluation of online financial content. From a practical perspective, universities, financial institutions, and policymakers in Kazakhstan may benefit from integrating digital financial literacy into educational initiatives for young adults. Such programs may help young investors make more informed decisions, reduce vulnerability to misleading financial information, and participate more responsibly in increasingly digital financial markets.

Despite these contributions, the findings should be interpreted in light of several limitations. The study focuses on young adults aged 18–29 and therefore may not fully represent other age groups. In addition, investment decision-making was assessed within a controlled research setting rather than through actual long-term investment behavior. Future studies may expand the sample, examine different demographic groups, and investigate whether improvements in digital financial literacy produce sustainable changes in real-world investment behavior over time.

References

Abdallah, W., Tfaily, F., & Harraf, A. (2025). The impact of digital financial literacy on financial behavior: Customers' perspective. *Competitiveness Review, 35*(2), 347–370. doi:10.1108/CR-11-2023-0297

Alhabsi, Z. S., Pandurengan, V., & Al Harthy, S. A. (2025). Evaluating the effect of financial literacy on investment decision: Meta-analysis. *Lex Localis – Journal of Local Self-Government, 23*(S6), 2951–2971. doi:10.52152/g35f7f76

Bayakhmetova, A., Bayakhmetova, L., & Bayakhmetova, L. (2023). Impact of financial literacy on investment decisions in developing countries: The example of Kazakhstan. *Asian Development Policy Review, 11*(3), 167–181. doi:10.55493/5008.v11i3.4917

Bhat, S. A., Lone, U. M., SivaKumar, A., & Krishna, U. M. G. (2025). Digital financial literacy and financial well-being: Evidence from India. *International Journal of Bank Marketing, 43*(3), 522–548. doi:10.1108/IJBM-05-2024-0320

Kumar, P., Islam, M. A., Pillai, R., & Sharif, T. (2023). Analysing the behavioural, psychological, and demographic determinants of financial decision making of household investors. *Heliyon, 9*(2), e13085. doi:10.1016/j.heliyon.2023.e13085

Ravikumar, T., Suresha, B., Prakash, N., Vazirani, K., & Krishna, T. A. (2022). Digital financial literacy among adults in India: Measurement and validation. *Cogent Economics & Finance, 10*(1), 2132631. doi:10.1080/23322039.2022.2132631

Sang, T. M. (2026). Integrating behavioral, social, and technological factors in cryptocurrency investment decisions. *PLOS ONE, 21*(4), e0332256. doi:10.1371/journal.pone.0332256

Shroff, S. J., Paliwal, U. L., & Dewasiri, N. J. (2024). Unraveling the impact of financial literacy on investment decisions in an emerging market. *Business Strategy & Development, 7*(1), e337. doi:10.1002/bsd2.337

Technical Sciences

SEISMIC DESIGN CONCEPT OF BUILDINGS AND SEISMIC RISK REDUCTION

Paata Rekvava
Dr. Prof., Georgian National Association for Earthquake Engineering and Engineering Seismology, 7, Lebanidze str. Tbilisi, Georgia

Abstract
The disasters, such as earthquakes that severely damage buildings, prove fatal to life. However, the truth remains that it is not the disasters that are responsible for the loss of lives, but the buildings that are incapable of withstanding these disasters. Loss life can be drastically reduced by ensuring that the structures of buildings are strong enough to bear the impact of earthquakes.
Buildings are a combination of art and science, which are put together to function as a living structure.
They are designed by architects and engineers. In reality, in most cases, buildings principally for human occupancy are designed conceptually by architects.
The three most important aspects of building design are: (i) Function, (ii) Form and (III) Strength. Function refers to the usage of the building and form refers to its overall geometry, composed in an aesthetically pleasing manner in keeping with the requirements of the user or intended functions. The form follows function. The architect is responsible for the form derived from the function of the building structure. The role of a structural engineer is to give enough strength his ideas and look after endurance and seismic strength of the structural members.
In Georgia, seismic resistant design is considered to be exclusively within the field of responsibility of structural engineers. As a result, architects are not well informed about the effect of their design decisions on the seismic performance of the buildings. Improving the collaboration and communication between
the architects and structural engineers can greatly help in reducing the seismic vulnerability of buildings. Successful collaboration can bring much more to a project than just facilitating its design and completion. The client and the users of a building are the primary beneficiaries of architect-structural engineer collaboration, but architects and structural engineers themselves benefit greatly.
The objective of this paper is to present the concept of seismic resistant design in a comprehensible format for the architects and structural engineers, therefore aims contribute to the efforts of creating a better awareness of seismic resistant building design.

Keywords: Seismic, Risk, Building, Conceptual, Design, Architect, Structural Engineer.

1. Introduction

The Georgian whole territory is one of the active part of Alpine-Himalayan earthquake belt. The tectonic structure of the region is determined by the location between the Eurasian and the Arabian plates and is characterized by a high level of seismic activity.
The tectonic rupture has a mosaic shape and there is no region without tectonic rupture. Javakheti's tableland is seismically active and the central part of the Caucasus is more active, such as Racha-Imereti border. It is noteworthy that tectonic rupture passes through the territory of

Tbilisi, the capital city of Georgia, what should be taken into account before constructing a high-rise buildings.

According to the current map of seismic hazard [1], the Georgian entire territory comprises 7(0.05-0.12g), 8(0.121-0.24g) and 9(0.241-0.52g) degree zones, where the earthquakes can be identified within the specified acceleration. Tbilisi is located in 8 degree zone (0.17g).Indicated quantities belongs to the grounds of average seismic properties. The design value of ground acceleration can be adjusted on the basis of seismic microzoning for the particular buildings site.

Earthquake engineering is integration of multidisciplinary knowledge which is engineering activity based on science and its goal represents to control seismic risk which are acceptable at socio-economic levels, human security and maintaining the material values. Its components are - engineering seismology, the theory of seismic resistance, design and construction practice.

Seismic risk can be defined as product of seismic hazard (intensity of event occurrence/probability) massive size of building and sensitivity towards earthquakes (about vulnerability/damages) and cost.

In the time of designing the buildings, the risk of an earthquakes is increasing because of the neglect of antiseismic construction principles and demands.

In terms of increased seismic risk, in order to get a socially safe, reliable building is require:

1. The project is to be developed according to the technical codes of earthquake engineering and a passed expertise by the staff of having an appropriate qualification and experience;
2. Quality building materials and the modern technologies should be used;
3. Building process must be conducted by professionals and the high quality of construction is necessary;
4. Author's supervision and quality control on construction site have to be organized.

According to the Georgian Government's Resolution N409 (18.06.2014) recognition of regulating the technical regulations of building sphere on Georgian territory and approval of operation, the amendment to the resolution N52 in 2014, January 14 temporarily allowed and recognized have to be the codes and rules of former Soviet Union, which was acting till 1992 year, and then was modified, other documents of technical regulations and those parts for the purpose of technical regulation of building sphere, alternative of which doesn't exist as a kind of adopted technical regulations or normative acts in Georgia, which aren't contrary to the applicable Georgian legislation or/and international treaties where Georgia is the participant.

2. Existing reality

Nowadays, in Georgia is a very difficult situation in construction sphere from the point view of seismic design standards. There is no united scientific-technical regulations of constructing-engineering activities. Intensive scientific-researches aren't being carried out in earthquake engineering sphere due to the lack of relevant profile institutions.

From 2010, National building standards are in De Jure action: earthquake engineering, concrete and reinforced concrete structures, foundations of buildings and structures and engineering climatology, which were approved in 2009 and of course it needs to be renewed in accordance with the modern standards and requirements. In the practice of design are also used:

European design standards -Eurocodes established by European Committee for Standardization (CEN);

International Building code (IBC), Section 1613 Earthquake Loads, established by International Code Council (ICC);

ISO 3010 - Basis for design of structures -seismic actions on structure, established by International Organization for Standardization (ISO).

In Georgian National Agency for Standards and Metrology there are registered:

609 EN standards in the sphere of construction (Civil engineering, building materials);

The En Eurocodes including 10 standards (EN1990-1999) (**Fig.1**) covering various subjects related to construction, are valid from 2009, May 15 as the cover methods of Georgian standards.

At the same time, today EN 0,1,2,3,7 and 8 Eurocodes are translated in Spatial Planning and construction Policy Department of Ministry of Economy and Sustainable Development of Georgia. Translated EN 1998 Eurocode 8 – Design of structures for earthquake resistance was discussed in the National Association for Earthquake Engineering and Engineering Seismology and corresponding comments were given to the Spatial Planning and Construction Policy Department.

National Annexes (NAs) of taking into consideration the Nationally Determined Parameters (NDRs) it is necessary to prepare by the local specialists.

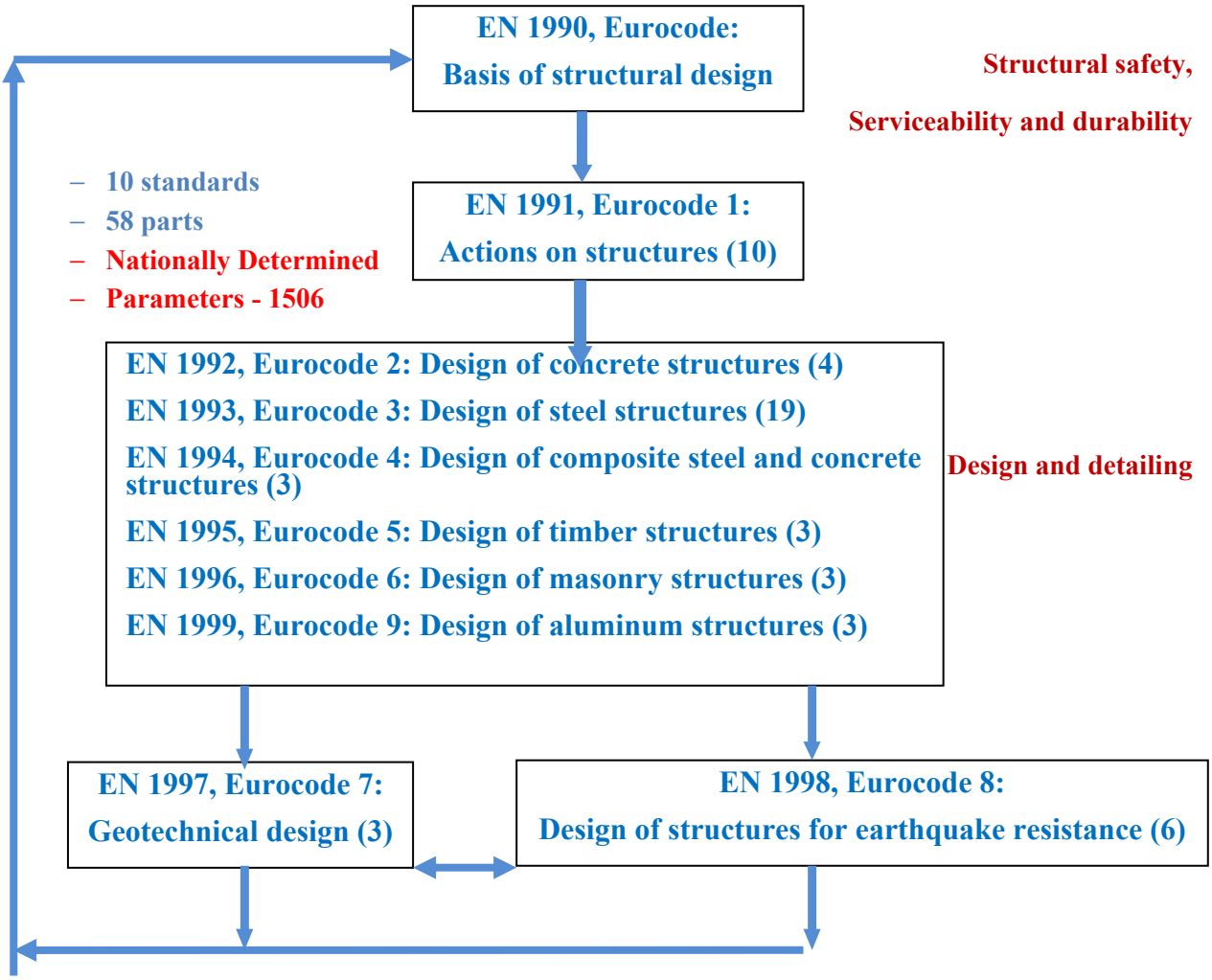


Fig.1. Interrelation of Eurocodes

Eurocode 8 EN 1998 – Design of structures for earthquake resistance (142 NDPs) consists of 6 parts:

EN 1998-1:2004 General rules, seismic actions and rules for buildings (61);
EN 1998-2:2005 Bridges (38);
EN 1998-3:2005 Assessment and retrofitting of buildings (12);
EN 1998-4:2006 Silos, tanks and pipelines (12);
EN 1998-5:2004 Foundations, retaining structures and geotechnical aspects (7);
EN 1998-6:2005 Towers, masts and chimneys (12).

In the event of earthquakes the objectives of EN 1998 are to ensure that:

- Human lives are protected;

- Damage of buildings is limited.

3. Modern ideology of earthquake engineering

According to Eurocode 8 constructions in seismic regions have to be designed and realized in the such way that

with the corresponding reliability to satisfy fundamental requirements:

No-life - threatening - collapse requirement-

Withstand the design seismic action without local or global collapse;

Retain structural integrity and residual load bearing capacity after the event.

This requirement should be met for a reference seismic action with PNCR=10% probability of exceedance in 50 years or with TNcR=475 years Return Period (corresponds to maximum design– a "rare "strong earthquake).

Damage limitation requirement-

Withstand a more frequent seismic action without damage;

Avoid limitation of use with high costs.

This requirement should be met for a reference seismic action with PDLR=10% probability of exceedance in 10 years or with TDLR=95 years Return Period (Corresponds to project – a" frequent" weak earthquake).

P and T values for a given country are defined in the National Annexes of this country standards.

Standard design is discussed in two ways in the theory of seismic resistance (Performance-based seismic design):

1. Ultimate limit state (ULS): Design of the structure for safety of people and/ or structure under a rare earthquake.

Plastic deformations and local damages are allowed in buildings.

2.Serviceability limit state (SLS) :concerns operation of the facility and damage to property under a frequent earthquake.

The damage limitation conditions must be fulfilled for buildings, i.e. the weak damages are allowed. The normal operational conditions and the appearance must be ensured.

Limit state (LS) design is the basis for all Eurocodes (including EC8).

One of the main provisions of Eurocode 8 is that aspect of seismic hazard in seismic regions must be considered at the initial stage of the conceptual design of the building, which allows the building to have constructive system and it responds to the basic requirements of seismic resistance while the expenses are acceptable.

Conceptual seismic resistant design is a subject to guiding principles and these principles are [3]:

Simplicity of building form, regularity in the plan (compact configuration and without divided contour) and height (constructive system without interruption or a certain intake);

Symmetry as in the plan also in a vertical projection (in terms of stiffness and mass distribution);

Uniformity and reservation of constructive elements (surplus of joining);

Resistance and stiffness with two main directions;

Torsional resistance and stiffness;

Rigid (diaphragmatic) working condition of intermediate floor;

Adequate foundation.

Conceptual designing is completed before the start of project document processing and clear constructive scheme of building within the frames of given parameters.

A building is a combination of art and science and three important aspects of its designing are:

1. Function (Destination);
2. Configuration (form, dimensions);
3. Stiffness.

American architect, Louis Sullivan stated in one of his poems [4], that the form follows function. This statement is now a doctrine for modern architects.

The architect is responsible for the function and configuration of the building. To give enough strength to the idea of architect and seismic resistance to constructive elements is the competence of structure engineer.

A tight cooperation between the architect and structural engineer has got a crucial importance during the phase of Conceptual designing.

4. Role of architect and structural engineer in the process of seismic resistant design

Architectural configuration plays a key role in seismic strength and functional conformity of the building.

Three main factors affecting on the configuration or concept of the building are as follows:

- Urban design, business (economics) and real estate issues;
- Planning and functional concerns, that impact the form of the building;
- Necessity and prestige of different image.

Urban design includes zoning and planning regulations.

The third influence, the need for the building to present an attractive, interesting, unique image to the outside observer and occupants.

Common problems of configuration and design mistakes what cause a damage or sometimes collapse of the building such as below:

- Flexible and not strong (weak) story;
- Variation of stiffness and strength in perimeter;
- Reentrant corners of buildings with the shape of L, T, X and H;
- Linked and short column;
- Strong beam - weak column;
- Mutual crashing – hammering effect of buildings;
- Non-regularity in the plan and height;
- Limiting height/ width ratio.

Analysis of literary sources [5, 6, 7, 8] shows that:

The building seismic resistance projecting are considered in many countries as the only sphere of structural engineers responsibilities and as a result, the architects are not well-informed about influence of design solutions on the seismic behavior of buildings;

As far as the building system is created by architect, therefore architects must be able to make assumptions in the process of conceptual designing about the behavior of building under seismic load.

Many architects maintain a wrong idea about structural engineer's involvement in designing work only at the time of strength calculation process. This is wrong opinion that can cause a significant result and additional cost.

None of calculation and detail constructing can compensate for the errors or mistakenly chosen unconstructive elements (frames filler walls, partitions, parapets, pediments etc.) and defects which were made in the process of conceptual seismic designing.

Architect:

Creates the project concept on the basis of projecting task (technical specification) worked out by customer;

Provides aesthetic and functional design of object;

Defines the type and efficiency of building system what can be used by structural engineer;

Coordinates the team of designers, provides the architectural part of project and technical interaction with other Neighboring parts;

Defines all project factors, which affect on the seismic behavior of building, such as:

- configuration of the building;
- regularity in height and a symmetry in the plan;
- type of structural framework system of the building;
- location - distribution of load-bearing constructive elements of each load such as columns and diaphragms;
- the type of unconstructive, architectural elements, dimensions and location-distribution;

Selects appropriate materials (high strength, ductility, homogeneity;
Provides "Seismic resistance architecture" – harmonization of architectural part of seismically optimized projects with requirements of the technical regulations of earthquake engineering and standards, ensures supervision.

Structural engineer:

Analyzes the technical specification worked out by the customer, participates in creation process of the building together with architect and carries out a constructive review of architectural solution of project object;
Works out a safe and effective constructive scheme of object in the seismic point of view;
Considers the types of materials for use in constructions and their properties, engineering-geological research results of the site;
Based on appropriate design model for constructive scheme of building organizes the proper calculation on forecast loads (static and seismic impacts) in accordance with the requirements of technical regulations of earthquake engineering;
Analyzes the calculation results and verifies them;
According to computed results realizes the detailing of constructive elements – determining the dimensions and reinforcing;
Provides correspondence of constructive part of project with the conceptual designing principles of seismic resistance and engineering supervision.

5. Collaboration between architect and structural engineer

Practice analysis of designing in seismic regions shows us that structural engineers are often very critical towards architects because of the facts that:
- They don't know the rules of anti-seismic design;
- They are working in a technical vacuum;
- Constructive prerogatives have been ignored;
- The flexibility and freedom of the design are put on front line;
- They don't try to get consultations at the initial stage for optimal constructive solutions;
- Made mistakes in the process of architectural designing, which are learned as a result of past Earthquakes is becoming the reason of building destruction in many cases;
- They demand from architects to improve the cooperation standards with them.

The architects point out that they are disappointed by structural engineers because of:
- They have innovation deficits;
- They have a bad contact with the idea of architectural design;
- They don't try to understand the demands and aspirations of architects.

Special literature indicates that general misunderstanding (incorrect understanding) between the architects, anti-seismic projecting rules and technical requirements those are restricting the creative freedom of designer.
Due to the differentiation of profession and role, dominance and methods of architects and constructors, their cooperation is a great challenge of two "Antagonistic" approach – Integration of the synthesis process characteristic for the architectural projecting and necessary design analysis for constructing.

The main position for successful interdisciplinary cooperation is understanding that it is more psychology than technology.
Many building projects require from architect and structural engineer to exchange information to themselves and work together what is very important:

- To meet the requirements of customer;
- For the compliance with proper standards;
- For the achievements of project coherent and safe exploitation of building.

Customers and building users are the main beneficiaries of cooperation with architect and structural engineer but the specialists also have the significant profit from the early synchronization of cooperation.

6. Recommendations for improving a collaboration

"Consecutive" or "Fragmented" design is worse and ineffective than the "Parallel" design. In the process of "Parallel" designing the architect and structural engineer got together to project a safe, effective, economic construction simultaneously designed on the static loads and seismic impact according to aesthetic and functional requirements in frameworks of given parameters (materials, configuration, height, constructive system, plasticity class, dissipation zones of seismic energy). Then they will choose unconstructive elements with allowed deformation for construction. By this approach, optimal designing is possible, when made choice becomes decisive for the seismic resistance and vulnerability of building.
This is the reason for why the architect and structural engineer should start working together on the project that seismic resistance engineering design could generate an aesthetic seismic resistance architecture.
For architects are recommended:
To start collaboration as soon as possible with structural engineer and other engineers for the better development of project at the stage of previous project/conception;
Harmonization of structural engineer's qualification and experience with project complexity;
Development of open and reliable culture at the time of project team leadership which will facilitate sharing of knowledge and encourage to approach the best consensus;
Developing a communication with the members of the project group so that everyone can make contribution and become responsible for the project at any stage;
To improve a knowledge about the issues of constructive and unconstructive seismic designs.
For structural engineers recommendations reduce to:
It is worth considering that architectural design includes a very wide range synthesis of different criterion and it is iterative and differ from constructive design;
As far as possible to encourage the project team and expecting that initial concept of project solution will need a development, perfection and correction in some cases;
Reviewing the generally acceptable solutions for innovation which will improve the project within the limits of maintaining the essence of engineer requirements and possible compromise;
Understanding the desired project architectural concept and quality before offering solutions.
Illustration of damages that are results of design mistakes is given on figures 2,3,4,5 below.

Fig.2 Strong Beam-Weak Column failures

Fig.3 Hammering Effect

Fig.4 Short Column effect and failure

Fig.5 Building with ground floor soft story

7. Conclusion

To sum up, we should realize that architectural configuration plays a major role in deciding the seismic strength and functional conformity of a building. The architect should shoulder the responsibility of ensuring proper seismic performance of the building since inception.
The building should be created/built to withstand earthquake forces rather then to retrofit after its completion. The project budget should be distributed wisely with the major share going into structural strength.
Improving the collaboration and communication between the architect, urban planning and earthquake engineering disciplines (seismology, structural engineering, lifelines engineering and emergency response) can greatly help in reducing the seismic vulnerability of buildings.

8. References

1. Building Code of Earthquake Engineering PN 01.01-09.Georgian Ministry of Economic Development, 2009.
2. The Eurocodes: Implementation and use. Joint Research Centre. European Commission.
3. Bachman H. Seismic Conceptual Design of Buildings – Basic principles of engineers, architects, building owners.
Swiss Agency for Development and cooperation. Biel 2002, 81p.
4. Charleson A. Seismic Design for Architects Outwitting the Quake, Elsevier, 2008.
5. Arnold C. Architectural Aspects of Seismic Resistant Design. Eleventh World Conference on Earthquake
Engineering. Paper No. 2003.
6. Charleson A.W. Pirie S. An Investigation of Structural Engineer-Architect Collaboration; SESOC Journal, new
Zealand, vol. 22, No1, April, 2009, pp.97-104.
7. Arslan M.M., Arslan M.D. An Architectural Evaluation of Structural Damages During Last Earthquakes in
Turkey. ECCOMAS Thematic Conference on Computational Method in Structural Dynamics and Earthquake
Engineering. Rethymno, Crete, Greece, 13-16 June, 2007
8. Rekvava P.A. Multi-factor assessment of seismic resistance of buildings and rehabilitation. TGU,GNAEEES, Tbilisi, 2022, 169p.

Medical Sciences

UDC: 616-006-071(574)

ORGANIZATION AND METHODOLOGY OF BREAST, CERVICAL AND COLORECTAL CANCER SCREENING: THE KAZAKHSTAN EXPERIENCE

Arman Khozhayev
Professor of the S.N. Nugmanov Department of Oncology, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Edgar Tagiev
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Tomiris Kemergazina
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Alikhan Mukhsanov
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Arailym Orazalina
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Valeriya Sidorova
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Diyar Suleyev
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Tamila Mussayeva
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Kazybek Amanturly
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Ayana Malikova
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
Arailym Sagynbek
Undergraduate Student, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan

Annotation: This scientific and analytical study examines the epidemiological patterns of cervical, breast, and colorectal cancer in the regions of the country, with particular emphasis on incidence and mortality rates and their regional variation. The study provides a comprehensive assessment of the current state of cancer screening and prevention, including the screening approaches used for the early detection of these malignancies in clinically asymptomatic populations. Particular attention is given to the organization and implementation of population-based screening programs, the target populations, screening intervals, diagnostic methods, and the subsequent management of individuals with abnormal screening results. The analysis also considers the role of screening in reducing the burden of malignant neoplasms through early detection and timely referral for further diagnostic evaluation and treatment. The study presents detailed, step-by-step algorithms for the organization and implementation of screening for cervical, breast, and colorectal cancer, outlining the principal stages from identification and invitation of eligible individuals to primary screening, additional diagnostic procedures, confirmation of the diagnosis, and appropriate follow-up. The diagnostic capabilities and practical significance of the screening methods currently applied are described, with particular emphasis on their potential for detecting precancerous conditions and early-stage malignancies before the onset of clinical symptoms. The findings highlight the importance of a systematic, evidence-based approach to cancer screening and demonstrate the potential of well-organized preventive programs to improve early detection, facilitate timely treatment, and ultimately contribute to reductions in cancer-related incidence and mortality.

Key words: oncology, primary care, early diagnosis, cancer screening, cervical cancer, breast cancer, colorectal cancer, epidemiology, incidence, mortality, HPV, Pap test, smear for oncocytology, ultrasound examination, mammography, hemocult test, fecal occult blood test - FOBT, total colonoscopy.

Early detection of malignant neoplasms remains one of the fundamental priorities of modern oncology and an essential component of effective cancer control. The clinical outcomes of many malignant diseases are strongly influenced by the stage at which the tumor is diagnosed, with earlier-stage detection generally associated with a greater range of therapeutic options, improved treatment outcomes, and more favorable long-term prognosis. Consequently, the development and implementation of effective strategies for identifying malignant and premalignant conditions before the appearance of clinical manifestations represent an important direction in contemporary preventive medicine. Cancer screening is an established approach to secondary prevention aimed at detecting disease in individuals who do not yet exhibit symptoms or clinical signs suggestive of malignancy. Unlike diagnostic procedures performed in patients with suspected disease, screening is conducted among predefined target populations according to established age, sex, and risk-based criteria. Its primary objective is to identify cancer at a preclinical or early stage, when the pathological process may be more amenable to curative treatment. In addition, screening may facilitate the detection of premalignant lesions and other pathological changes that have the potential to progress to invasive cancer, thereby creating an opportunity for timely intervention and prevention of malignant transformation [1].

The effectiveness of a screening program depends not only on the diagnostic performance of individual screening methods but also on the quality of its overall organization and implementation. A well-functioning screening system requires clearly defined target populations, standardized procedures for participant recruitment and examination, appropriately trained healthcare professionals, and consistent criteria for interpreting and managing screening results. Equally important is the availability of an established pathway for the subsequent evaluation of individuals with abnormal findings, including confirmatory diagnostic procedures, specialist

referral, treatment, and follow-up. Without adequate coordination between these stages, the potential benefits of screening may be substantially reduced [1,2].

Screening technologies and diagnostic approaches should meet several fundamental requirements to ensure their applicability at the population level. Ideally, screening procedures should be sufficiently simple, safe, accessible, reproducible, and acceptable to the target population, while demonstrating adequate sensitivity and specificity. High sensitivity is essential for minimizing the number of clinically significant cases that remain undetected, whereas adequate specificity helps to reduce false-positive results and unnecessary diagnostic interventions. Therefore, the development of standardized, evidence-based screening algorithms and their systematic implementation are critical for improving the early detection of cancer and precancerous conditions and, ultimately, for reducing cancer-related incidence and mortality [3].

Screening plays an important role in improving early diagnosis and treatment outcomes. According to the Guide to Cancer Early Diagnosis by Ilbawi A. et al. [4], screening aims to detect unrecognized cancer or its prior lesions in a typically healthy, asymptomatic population through tests or other procedures that can be applied quickly and are widely available to the target population. In screening, the target population is assessed for unrecognized cancer or precancer, and most people tested will not be diagnosed with the disease. Screening should be seen as a process and not as the performance of a specific test, examination, or procedure. The screening process includes a system of informing and inviting the target population to participate; administering the screening test; following-up with test results and referral for further testing among those with abnormal test results; ensuring timely pathologic diagnosis, staging and access to effective treatment with routine evaluation to improve the process. A screening program encompasses the process from invitation to treatment and requires planning, coordination and monitoring and evaluation.

To date, the republican cancer screening program includes three nosological forms of malignant neoplasms - cervical cancer (CC), breast cancer (BC), colorectal cancer (CRC). Let's consider the current epidemiological indicators, methodology and results of cancer screening in our country.

CC in the structure of all malignant tumors of both sexes of the population in 2022 took 6th place with a share of 5.51% (2021 - 4th place, 5.54%), in women - stable 2nd place - 9.7% (9.7%) [5].

The incidence rate per 100 thousand population increased from 9.4 to 9.92. In 10 regions of the republic, the incidence rate is higher than the national average: Pavlodar - 17.2 per 100 thousand people (2021 – 16.7) – the highest level, East Kazakhstan – 14.3 (10.8), North Kazakhstan – 14.3 (10.2), Atyrau – 13.2 (13.8), Zhetysu - 11.7, Karaganda - 11.7 (12.0), Abay - 11.1, Akmola - 11.1 (11.9), Mangistau - 11.1 (9.7), Kostanay - 10.8 (10.6) regions.

Low incidence rates in Zhambyl region - 5.8 per 100 thousand population (5.7), Turkestan region - 6.1 (5.2), Aktobe region - 8.3 (11.6), Kyzylorda region - 8.5 (8.2) areas.

CC in the structure of causes of death from malignant tumors of the population of both sexes in 2022 rose from 9th to 8th position, with a share of 4.6% (2021 - 4.3%), mortality from CC is stable at 3.1 per 100 thousand population (3.1).

The mortality rate from CC in 10 regions is higher than the national average: Akmola - 4.2 per 100 thousand population (2021 - 3.1) - maximum level, West Kazakhstan - 4.1 (4.8), Pavlodar - 3.8 (5.6), Almaty – 3.7 (2.5), Zhetysu – 3.7, Atyrau – 3.4 (4.0), East Kazakhstan – 3.3 (3.8), Karaganda - 3.2 (4.7), Kostanay - 3.2 (2.4) regions and Almaty city - 3.4 (2.9).

Below the national average, mortality was recorded in Abay region, cities Astana, Shymkent - 2.9 per 100 thousand population, Mangistau - 2.8 (3.0), Turkestan - 2.3 (2.2), Aktobe - 2.2 (3.0), North - Kazakhstan - 2.0 (2.6), Kyzylorda regions - 1.7 (3.5) - the best result [5].

In 12 regions, a 100% level of morphological verification of the diagnosis was ensured, the

lowest or worst indicator for the third year was in the Kyzylorda region - 94.3%, below the national average indicators in Akmola - 98.8%, Atyrau - 98.9%, Kostanay - 98, 9%, Mangistau - 97.6%, Pavlodar - 96.6%, regions and Almaty city - 98.5%;

In a number of regions, the frequency of diagnosis of stage I-II CC was below the national average (88.1%) - in Akmola - 76.2% (2021 - 73.6%) - the worst result in the country, in Karaganda - 77, 2%, Zhetysu - 82.9%, Abay - 83.8%, Kostanay - 84.3%, Aktobe - 85.5%, West Kazakhstan - 85.7%, Pavlodar - 81.3%, while that in the Atyrau region - 100.0% result.

The proportion of stage IV CC is higher than the national average (2.7%) in the following regions: the worst result is in Zhetysu (6.1%), above the national average in Karaganda - 5.1% (2021 - 5.6%), Akmola - 4.8% (2.3%), Kostanay - 4.5% (4.4%), North Kazakhstan - 3.9% (7.4%), Almaty - 3.7 % (5.1%), Zhambyl - 2.9% (0.0) regions, cities Almaty – 3.6% (1.8%) and Shymkent – 3.8% (5.9%). The lowest neglect is in the East Kazakhstan region - 1.0% (0.7%).

Late diagnosis rates (III-IV stages) for CC are above the national average - 11.9% (15.4% in 2021) were noted in Akmola - 23.8% (2021 - 26.4%) - worst result, Karaganda - 22.8% (35.2%), Pavlodar - 18.8% (20.8%), Zhetysu - 17.1% (24.2%), Abay - 16.2% (12.8%), Kostanay - 14.6% (15.6%), Aktobe - 14.5% (9.6%), West Kazakhstan - 14.3% (32.4%) regions. The lowest neglect is in the Mangistau region - 6.0% (20.8%).

Across the country, the five-year survival rate of patients with CC registered in 2018 was 59.9% in 2022, with a decrease from the level of 2021 (67.5% for those registered in 2017), and with a significant range in by region, from the maximum – 72.9% (2021 – 70.7%) in the North Kazakhstan region, to the minimum – 34.9% (64.4%) in the Atyrau region [5].

CC screening is a periodic, comprehensive examination of women of a certain age group as part of a special medical program to prevent and reduce incidence and mortality from CC [6].

CC screening is a population-based preventive intervention aimed at detecting cervical precancerous lesions and CC at an early stage, when timely diagnostic assessment and treatment can prevent progression to invasive disease. In Kazakhstan, CC screening is organized within the national population screening program and is conducted among predefined target groups according to age and clinical eligibility criteria [7,8,9].

According to the current national screening framework, women who reach the ages of 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70 and 74 years and who are not under dynamic follow-up for cervical malignancy constitute the target population for CC screening. Screening is performed once every four years, with the screening episode expected to be completed within 60 days.

The age-specific schedule represents a population-based strategy in which eligible women are identified through the primary healthcare system. Recent data from Kazakhstan confirm the use of a four-year screening interval and cytological examination for CC screening within the national program.

The preparatory stage is performed within the primary healthcare organization. It includes:

- identification and annual formation of the target population;
- verification and monthly updating of the list of eligible women;
- informing women about the purpose and benefits of CC screening;
- invitation to participate in screening;
- registration of screening attendance and organization of timely completion of the screening pathway.

The national regulations require primary healthcare organizations to form target groups from the population attached to the healthcare organization and to establish a sequential pathway for the provision of screening services. The target population is initially compiled annually, with subsequent monthly correction.

The primary screening examination is performed in the primary healthcare setting by an

appropriately trained healthcare professional. The screening procedure includes collection of a cervical specimen for cytological examination.

Under the current Kazakhstan program, cervical screening is based on cytological examination of a cervical smear using the Papanicolaou (Pap) test, with liquid-based cytology used within the national screening program. The collected specimen is submitted to a cytology laboratory for standardized evaluation.

Before specimen collection, the woman should receive appropriate instructions concerning factors that may interfere with specimen quality. The collected material must be properly identified, documented and transported according to laboratory requirements.

Cytological examination should be reported using standardized terminology, preferably the Bethesda System, which provides a reproducible framework for classifying cervical cytological findings [10].

The cytological result may be classified as:

- NILM (Negative for Intraepithelial Lesion or Malignancy) – no evidence of intraepithelial lesion or malignancy;
- ASC-US (Atypical Squamous Cells of Undetermined Significance);
- ASC-H (Atypical Squamous Cells, cannot exclude HSIL);
- LSIL (Low-Grade Squamous Intraepithelial Lesion);
- HSIL (High-Grade Squamous Intraepithelial Lesion);
- AGC (Atypical Glandular Cells) and other glandular abnormalities;
- cytological findings suspicious for or diagnostic of invasive carcinoma.

The Bethesda terminology is widely used in contemporary cervical screening and abnormal-result management. Current evidence-based guidelines emphasize that management should depend not only on the cytological category but also, where available, on Human papillomavirus (HPV) status, previous screening history and the estimated risk of clinically significant cervical intraepithelial neoplasia.

A woman with an abnormal cytological result should not simply return to routine screening. She should enter a diagnostic pathway according to the severity of the abnormality and the applicable national clinical recommendations.

The diagnostic pathway may include: abnormal cytology → clinical assessment → HPV testing where indicated → colposcopy → directed cervical biopsy when indicated → histological examination → definitive diagnosis → treatment or surveillance.

Risk-based management is preferable to management based exclusively on a single cytological result. The 2019 American Society for Colposcopy and Cervical Pathology consensus guidelines, published in 2020 and subsequently updated, incorporate current screening results together with previous results and clinical history when determining the need for surveillance, colposcopy or treatment.

For example, high-grade cytological abnormalities generally require prompt diagnostic evaluation, while selected low-grade abnormalities may be managed with HPV-based triage and/or surveillance depending on the patient's estimated risk. Unsatisfactory cytology requires repeat testing or further evaluation according to the applicable guideline.

The second stage of the national screening pathway is intended to determine the cause of an abnormal screening result and establish a definitive diagnosis. Under the Kazakhstan regulations, women with abnormal screening findings are referred by the primary healthcare physician for further examination. Depending on the screening result, this may include consultation with an obstetrician-gynecologist, colposcopy or videocolposcopy, cervical biopsy when indicated, histological examination and consultation with an oncologist or gynecologic oncologist.

Thus, the screening process should be regarded as a continuum extending from population

identification and primary testing to diagnostic confirmation and subsequent treatment or surveillance, rather than as an isolated cytological examination.

At the final stage, the screening results and the established clinical diagnosis are documented in the healthcare information system. The woman is informed of the results and receives recommendations regarding further diagnostic evaluation, treatment, surveillance or return to routine screening.

The national regulations specifically require screening organizations to record screening results, final diagnoses and recommendations for further examination and follow-up. The screening episode is considered complete when the appropriate endpoint of the clinical pathway has been documented.

The contemporary population-based cervical cancer screening algorithm can therefore be summarized as follows:

Identification of the target population → invitation and information → cervical specimen collection → primary cytological screening → cytological classification → assessment of abnormal results → HPV testing/triage where indicated → colposcopy and biopsy when indicated → histological confirmation → treatment or surveillance → documentation and return to routine screening.

This approach combines the organizational framework of the national screening program in Kazakhstan with contemporary principles of risk-based management of abnormal cervical cancer screening results.

Now, regarding the results of CC screening. In 2022, 771,282 women of the target group aged 30 to 70 years were examined during cytological screening (in 2021 - 757,454).

During cytological screening in 2022, 392 cases of cervical cancer were identified (319 in 2021). The detection rate increased from 0.42 to 0.51 per 1000 women examined

High detection of CC during screening is ensured in Aktobe, Almaty, Atyrau, East Kazakhstan, Kyzylorda, Pavlodar, North Kazakhstan, Turkestan regions and Shymkent city. The detection rate in these regions ranges from 0.55 to 1.59 per 1000 women examined. The best indicator is in Atyrau region - 1.59. Compared to 2021, there is an increase in detection in 10 regions, with the exception of Akmola, Aktobe, Zhambyl, Kostanay, Mangistau, North Kazakhstan regions and Shymkent city. The worst result in Astana is 0.15 per 1000 women examined [5].

Cytologically, cervical precancer was detected in 1.16% of those examined (2021 – 0.99%). The detection rate of precancer below 0.6% (the planned indicator for 2022, according to the Comprehensive Plan) was noted in Aktobe, Karaganda and Kostanay regions.

A high proportion of stage I CC (70% or more) was detected in 6 regions of the country (in 8 in 2021): Kostanay, Mangistau (94.7% - best result), North Kazakhstan, Turkestan regions, cities Almaty and Astana. Low levels of early detection of CC (below 50%) were not observed in any region.

Localized processes (stages I-II) were identified in 99.2% of all cases of detected cancer (96.5%). In the Akmola and Karaganda regions, cases of CC were identified not only in localized, but also in widespread stages of the process. A total of 3 cases of CC in stage III and no cases in stage IV were identified (11 and 0, respectively) [5].

BC ranks first in the structure of the frequency of malignant tumors of both sexes in the population with a share of 14.7% (2021 - 15.4%). This situation has been stable since 2004; in addition, BC ranks first and remains consistently in this position in the structure of female oncopathology. The incidence of BC in 2022 in the country as a whole increased to 26.5 per 100 thousand (2021 – 26.3). In the structure of cases, BC occupies the 1st ranking place in the vast majority of regions and cities of the country, except for three: Akmola, Kyzylorda and North Kazakhstan regions, where lung cancer takes the 1st ranking place [4].

Above the national average - 26.5 per 100 thousand of us. – incidence of BC in 10 regions

of the country: Abay – 33.3, Akmola – 32.7 (2021 – 29.8), East Kazakhstan – 44.7 (39.9) – the highest level, West Kazakhstan – 31.2 (28.4), Karaganda – 40.2 (40.1), Kostanay – 37.5 (35.8), Pavlodar – 43.2 (47.4), North Kazakhstan – 34.7 (38.2) regions and Almaty city – 35.4 (34.5), Astana city – 31.5 (28.4). Below average indicators per 100 thousand of us. in Aktobe - 21.6 (24.3), Almaty - 21.9 (17.7), Atyrau - 22.8 (15.7), Zhambyl - 14.2 (15.1), Zhetysu - 22.8, Kyzylorda - 14.6 (14.4), Mangistau - 14.7 (17.3), Turkestan - 11.3 (11.7) regions and Shymkent city - 14.9 (21.9) [5].

BC ranks third in the structure of causes of death from malignant tumors in the population of both sexes for the thirteenth year in a row, amounting to 8.1% in 2022 (2021 – 8.7%). In the republic as a whole, mortality from BC decreased by 13.0%, from 6.2 to 5.4 per 100 thousand people.

The regions where mortality from BC is higher than the national average include: Abay - 10.1 per 100 thousand people (maximum level), East Kazakhstan - 8.0 (2021 - 8.5), Pavlodar - 7.1 (10.0), North Kazakhstan - 7.0 (11.4), Kostanay - 6.9 (7.5), Akmola - 6.5 (8.2), West Kazakhstan - 5.7 (6.9), Zhambyl - 5.5 (4.8) and Astana city – 6.3 (6.6), Almaty city – 6.6 (9.5). The indicators are significantly lower in Aktobe - 4.5 (3.5), Almaty - 4.5 (5.8), Zhetysu - 4.0, Atyrau - 3.7 (3.0), Kyzylorda - 4.4 (4.1), Turkestan - 3.6 (3.6), Mangystau regions - 2.7 (3.6) - the lowest level [5].

Mass screening to identify BC patients should mainly involve healthy women without any signs of the disease or symptoms. Screening not only helps to detect hidden forms of cancer that can be treated, but also has psychological value for women. As a result of screening, women are convinced that they do not have BC, and this is the most important potential success of such programs. While the ultimate goal of screening is to reduce BC mortality, its immediate goal is to detect cancer before clinical manifestation. However, BC is a heterogeneous disease, which can significantly affect the effectiveness of screening. Screening models for BC are usually based on the fact that the majority of detected tumors are invasive cancers in the early stage of progression. In addition, it must be taken into account that the detection of cancer (or its precursors) before clinical manifestation increases the risk of false positive diagnosis [11,12].

Mammography has a sensitivity of 95% and a specificity of 97%. These indicators decrease when examining women with denser mammary glands (young age, use of hormone therapy), with low quality mammography, and also with insufficient qualifications of the radiologist. Detection of high-grade invasive cancer by screening, when the tumor is not yet detected by clinical examination (palpation), means the possibility of reducing mortality from BC [13].

Preventive screening for early detection of BC in the Republic of Kazakhstan includes [7]:

1) mammography of both mammary glands in two projections - direct and oblique in the mammography room of the city, district polyclinic (mobile medical complex). All digital mammograms in the presence of a system for archiving and transferring medical images are copied to CDs and other electronic media and transferred to the server of the mammography room of the Cancer Center using specialized licensed software integrated between medical organizations; in case of impossibility of digital transmission - they are printed on X-ray film at a scale of 1:1 - 100% (1 patient - 1 set - 2 or 4 mammograms) with subsequent transfer to the mammography room of the Cancer Center;

2) interpretation of mammograms according to the BI-RADS classification (M0t, M0d, M1, M2, M3, M4, M5) by two or more independent radiologists of the same medical organization - double reading or different medical organizations: a radiologist of the mammography room city, district polyclinic (mobile medical complex) - the first reading, and the radiologist of the mammography room of the Cancer Center - the second reading;

3) in-depth diagnostics - targeted mammography, ultrasound examination (hereinafter - ultrasound) of the mammary glands, trepanobiopsy, including under ultrasound or stereotaxic control for histological examination, which is carried out in case of detection of pathological changes on mammograms (M0d) in the mammography room of the Cancer Center.

√ An average medical worker or a responsible person of the organization of outpatient care sends the patient for mammography to the district, city polyclinic.

√ The X-ray laboratory assistant of the mammography room of the city, district polyclinic (mobile medical complex) performs mammography, fills out a referral for double reading of mammograms and transmits the referral through information interaction.

Radiologist of the mammography office of the city, district polyclinic (mobile medical complex): fulfills the requirements for the safety and quality of mammographic examinations; evaluates the quality of the images provided and the correctness of the installation; performs repeated mammography in the M0t category (technical errors of mammography); determines the radiological density of the mammary glands on the ACR scale (A, B, C, D) indicating this parameter in the study protocol; conducts the first reading of mammograms with interpretation of the BI-RADS classification results. In the M0d category (undetermined or suspicious radiological changes requiring additional examination), the study protocol indicates the predominant pathology: education, asymmetry, violation of architectonics, microcalcifications; sends mammograms, electronic copies of mammograms through the archiving system and transfer of medical images to the workplace of the mammography office of the Cancer Center together with directions for double reading of mammograms; directs low–dose computed tomographic images through the system of archiving and transferring medical images to the workplace of the computer tomography office of the Cancer Center together with copies of images recorded on CD-ROMs or other electronic media and directions for double reading.

♦ The radiologist of the mammography room of the Cancer Center: evaluates the quality of the provided images and the correctness of the styling. Viewing digital x-ray images transferred to the server or on digital media (CD, DVD) is carried out on a monitor for interpreting digital x-ray images with a resolution of at least 5 megapixels, which has a certified grayscale transmission in accordance with the DICOM standard; conducts a double (second) reading of mammograms with the interpretation of the results according to the BI-RADS classification, using, if necessary, archival images. Organizes the third reading according to indications. With double reading, an independent interpretation of the images is carried out (blinding method - the second radiologist does not know the results of the first reading); in the M0m category (technical errors in mammography), recommends repeat mammography; in the M0d category (uncertain or suspicious radiographic changes requiring additional examination), the study protocol indicates the predominant pathology: education; asymmetry, violation of architectonics, microcalcifications; recommends that the outpatient care organization, according to indications, invite the patient for in-depth diagnostics (targeted mammography, ultrasound of the mammary glands, trephine biopsy, including under ultrasound or stereotaxic control, followed by histological examination of the material); collects and archives all mammograms (films and electronic media) made as part of the examination. The shelf life of mammograms is at least 3 years after leaving the age subject to a screening study; the results of the double (second) reading are transferred to the outpatient care organizations through information exchange.

♦ Indications for in-depth diagnostics are the conclusions of double reading mammograms M0d (uncertain or suspicious X-ray changes requiring additional examination).

♦ In-depth diagnostics is carried out in two stages. At the first stage, ultrasound is performed, according to indications, targeted mammography, possibly with an increase (with asymmetry, violation of architectonics and the presence of microcalcifications). When visualizing a suspicious pathology (M4 and M5), the second stage is performed - trepanbiopsy, including under ultrasound control and stereotaxic control for histological examination.

♦ Histological examination is carried out in the laboratory of pathomorphology or pathological bureau. Morphological interpretation of the biopsy is carried out in accordance with the recommendations of the World Health Organization.

♦ Physician or responsible person of the outpatient care organization:

1) upon receipt of a mammography result according to the BI-RADS classification:

- in case of M0t (technical errors in mammography) - sends the patient for a second X-ray examination to the mammography room of the city, district polyclinic (mobile medical complex);

- with M0d (undefined or suspicious X-ray changes requiring additional examination) - sends the patient for in-depth diagnostics to the mammography room of the Cancer Center;

- with M1 (no changes detected) - recommends that the patient undergo a follow-up mammography examination after 2 years. With radiological density of the mammary glands, C and D are sent for ultrasound of the mammary glands to exclude a false-negative result of mammography;

- with M2 (benign changes), refer the patient for a consultation with an oncologist (mammologist) of the clinical diagnostic department, followed by a screening mammography examination after 2 years;

- with M3 (probable benign changes) - sends the patient for short-term dynamic radiation observation to the local doctor with the recommendation of control mammography or ultrasound in 6 months;

- with M4 (signs that cause suspicion of malignancy), M5 (practically reliable signs of malignancy) and if it is technically impossible to perform a trepanbiopsy or a biopsy is refused, a referral to an oncologist (mammologist) of the clinical diagnostic department for dynamic observation and decision on the verification of the identified pathology;

2) upon receipt of the result of a histological examination:

- benign education - refers the patient to an oncologist (mammologist) of the clinical diagnostic department for dynamic monitoring, followed by a screening mammography examination after 2 years;

- formation with an indeterminate malignant potential or carcinoma in situ - refers the patient to the Cancer Center for consultation and treatment, followed by dynamic observation by an oncologist (mammologist) of the clinical diagnostic department at the place of her attachment;

- malignant neoplasm - refers the patient to the Cancer Center for treatment and follow-up;

3) communicates the results of the screening examination to the patient in any available way (by telephone, in writing, through electronic means of communication);

4) enters the results of double reading, in-depth diagnostics, histological examination, recommendations of the radiologist of the Cancer Center mammography room into the information system.

Establishing the size of the primary tumor is especially important in screening. Tumor size is an important criterion for evaluating the quality of screening and determining the ability of X-ray mammography to detect non-palpable tumors. Therefore, it is extremely important that pathologists measure tumor diameter as accurately as possible. The smaller the size of the primary tumor, the greater the likelihood of error in determining its size.

Let's analyze the results of BC screening. Mammography screening identified 1,570 cases of BC in 2022 (1,402 in 2021). The cancer detection rate increased from 1.78 to 1.94 per 1000 examined. The best result is in the Karaganda region – 2.63 per 1000 women examined. Low detection rate per 1000 examined, compared to the republican average, in Atyrau (1.72), Zhambyl (0.58), Kyzylorda (1.68), Mangistau (0.42 - worst result), Turkestan (1.22) regions and cities Astana (1.5) and Shymkent (1.58). Compared to 2021, there was an increase in the detection of BC in 9 regions, with the exception of Aktobe (decrease from 2.87 to 2.19 per 1000 women examined), Karaganda (from 2.73 to 2.63), Mangistau (from 1.10 to 0.42), North Kazakhstan (from 3.27 to 2.31), Turkestan (from 1.36 to 1.22) regions and cities Astana (from 1.54 to 1.50), Almaty (from 2.24 to 2.18) and Shymkent (from 2.35 to 1.58) [5].

In 2022, the proportion of patients identified during screening studies with early stages of BC (stage 0-I) was 50.2% during screening (in 2021 - 47.9%). A high proportion of stages 0-I BC (over 50%) was recorded in 8 regions (in 8 in 2021): Akmola, West Kazakhstan, Karaganda (70.8% - best result), Pavlodar, North Kazakhstan, Turkestan regions, cities Astana and Shymkent. Low levels of early detection of BC (below 40%) were noted in Aktobe (19.3% - worst result), Zhambyl (34.8%), Kostanay (39.5%), Mangistau (27.3%) regions and Almaty city (37.3%). Localized cancer (0-I and II stages) amounted to 96.2% (2021 - 95.5%), while not a single case was detected in stages III-IV in Atyrau, West Kazakhstan, Zhambyl, Kyzylorda, Mangistau, Pavlodar regions, cities Astana and Shymkent. A total of 46 cases of breast cancer in stage III and 14 in stage IV were identified (52 and 11, respectively) [5].

Epidemiological indicators of CRC in the form of colon cancer and colorectal cancer are considered separately for objective reasons.

Colon cancer with a specific gravity of 5.53% (2021 - 5.2%) in the structure of oncopathology of both sexes of the population has risen to 5th place, in men it remains in 6th place - 5.8% (5.5 %), for women - in the 5th - 5.3% (4.91%) The incidence rate of cancer of this localization in the country in the reporting year increased from 8.8 to 9.95 per 100 thousand population.

The incidence of colon cancer in 10 regions is higher than the national average - 9.95 per 100 thousand population: Kostanay - 20.7 (2021 - 15.9), Pavlodar - 18.8 (15.3), North Kazakhstan - 18, 0 (12.7), East Kazakhstan - 16.9 (13.4), Karaganda - 15.4 (15.0), Akmola - 14.6 (10.2), West Kazakhstan - 11.0 (10.1), Abay - 10.0 (9.0) regions and cities Almaty – 12.8 (12.1) and Astana – 10.5 (9.0). As in 2021, colon cancer was detected much less frequently in Turkestan - 3.1 per 100 thousand population (2.7), Kyzylorda - 4.1 (4.6), Zhambyl - 5.5 (5.8), Almaty - 6.3 (4.7), Zhetysu - 6.4, Mangistau - 6.8 (4.9) regions and Shymkent city - 5.0 (4.0) [5].

Rectal cancer in the structure of malignant neoplasms of both sexes retains 7th place in rank with a specific gravity of 4.9% (2021 - 4.92%), but in men it dropped from 4th to 5th place - 6.1%, for women – from 9th to 10th – 4.0%. The incidence rate per 100 thousand population increased from 8.4 to 8.8.

A high incidence rate was recorded in Kostanay - 17.8 per 100 thousand population (2021 - 16.2), East Kazakhstan - 17.7 (13.9), North Kazakhstan - 15.6 (15.1), Pavlodar – 14.9 (18.1), Karaganda – 13.3 (11.7), Abay – 12.9, West Kazakhstan – 12.9 (9.8), Akmola – 10.3 (13.1) regions and Astana city – 10.3 (9.0). Traditionally, a low incidence of rectal cancer is observed in Mangistau - 3.1 (2.8), Turkestan - 3.3 per 100 thousand population (2.7), Zhambyl - 3.7 (5.1), Kyzylorda - 4, 1 (5.3), Almaty – 5.3 (5.6) regions and in Shymkent city – 5.5 (5.0) [5].

Rectal cancer in the structure of causes of death from malignant neoplasms of the population of both sexes in 2022 remained in 5th place with a share of 5.41% (2021 – 5.41%). In the republic as a whole, the mortality rate from this form of cancer was 3.6 per 100 thousand population (3.87).

The mortality rate per 100 thousand population was higher than the national average in East Kazakhstan - 7.8 (2021 - 8.6) - the maximum level, Pavlodar - 7.5 (7.6), Abay - 5.9, North Kazakhstan - 5.8 (4.3), Kostanay - 4.9 (4.9), West Kazakhstan - 4.8 (4.2), Karaganda - 3.8 (5.2) regions. Below the national average - 3.8 per 100 thousand population, mortality in Aktobe - 3.2 (4.1), Almaty - 2.6 (2.6), Atyrau - 2.5 (3.4), Zhetysu - 2, 6, Zhambyl - 3.3 (2.7), Turkestan - 2.1 (1.6), Mangistau - 1.9 (1.2), Kyzylorda regions - 1.8 (2.1) - the lowest figure , and cities Almaty – 3.7 (4.3), Shymkent – 2.6 (2.1).

Colon cancer in the structure of causes of death from malignant neoplasms of the population of both sexes in 2022, as in 2021, ranks 6th, with a share of 5.2% (2021 – 5.0%). At the same time, the mortality rate in the country decreased by 5.6%, from 3.6 to 3.4 per 100 thousand population.

Mortality rates in 10 regions are higher than the national average: East Kazakhstan - 7.1 per 100 thousand population (2021 - 5.1) - maximum level, Pavlodar - 5.6 (6.0), Kostanay - 5.3 (5.6), Akmola – 5.2 (3.8), Abay – 5.1, Karaganda – 5.1 (5.6), West Kazakhstan – 4.8 (4.4), North Kazakhstan – 4.8 (5.0) regions and cities Astana – 3.6 (2.7), Almaty – 4.5 (5.3). Low mortality rates from colon cancer were noted in Kyzylorda - 1.2 per 100 thousand population (2.7) - the best result, Turkestan - 1.3 (1.7), Mangistau - 1.6 (2.6), Aktobe – 2.0 (2.5), Zhetysu – 2.4, Zhambyl – 2.5 (3.7), Atyrau – 2.5 (1.8), Almaty – 2.6 (1.8) regions and cities Astana – (2.7), Shymkent – (2.4).

For colon cancer (94.0%) - 100% verification level was achieved in 3 regions (Abay, Almaty and Turkestan regions), high rates in the Astana city (98.5%), Shymkent city (98.0%), Zhambyl (98.4%), Atyrau (98.2%) regions, low – in Akmola region (86.7%), Almaty city (84.3%), in the Kyzylorda region (61.8%) – the worst result since 2017.

For rectal cancer (97.4%) - in 6 regions there is a 100% verification level, the worst level is still in the Kyzylorda region - 85.3%, lower than the republican average in the Akmola region - 92.6%, Aktobe region - 96 .8%, Mangystau region - 87.0%, Pavlodar region - 95.3%, Almaty city - 93.2% [5].

The frequency of diagnosis of stage I-II rectal cancer, as a visually accessible localization (68.9% - national average) in the regions, was: in Akmola - 34.6% - the worst result, as in 2021, in the country (2021 - 44.1%), Mangistau - 47.8%, Abay - 53.9%, West Kazakhstan - 59.1%, Almaty - 66.2%, Zhetysu - 68.6%, Karaganda - 65, 7% regions and Shymkent city - 62.9%.

For colon cancer (52.4%), early diagnosis rates are higher in Pavlodar (65.9% - best result), Abay, Aktobe, Atyrau, East Kazakhstan, Zhambyl, Zhetysu, Karaganda, Kostanay, Pavlodar, North Kazakhstan, Turkestan regions and Shymkent. The lowest figure (23.5%) is in the Kyzylorda region.

For colon cancer (17.3%), the rates of neglect at stage IV are higher - in Akmola - 31.0% - the worst result (2021 - 20.3%), Zhetysu - 27.3%, Abay - 23.1% , Turkestan - 22.2% (29.1%), Karaganda - 28.1% (28.6%), West Kazakhstan - 18.8% (8.2%), Mangistau - 17.6% (19 .4%) regions and cities Astana - 18.0% (22.9%), Shymkent - 20.0% (22.7%). The lowest level of neglect is 2.9% in the Kyzylorda region (7.9%).

The proportion of stage IV in rectal cancer (13.1%) is higher in Akmola - 29.6% - the worst result (2021 - 19.4%), Abay - 19.7%, Kyzylorda - 17.6% (9.1%), Karaganda - 16.9% (28.4%), Almaty - 15.6% (17.0%), Kostanay - 14.8% (11.1%), Zhambyl - 13.3 % (13.6%) regions and Shymkent city - 14.5% (12.5%). The lowest level of neglect - 6.0% - is in the Atyrau region (12.5%).

Late diagnosis of rectal cancer as a visually accessible localization (stages III-IV) in 2022 amounted to 31.1% (in 2021 - 33.5%).

For rectal cancer, the level of neglect is higher than the national average - 31.1%, the indicators in Akmola - 65.4% (2021 - 55.9%) - the worst result in the country, Mangistau - 52.2% (38.1%), Abay – 46.1% (30.6%), West Kazakhstan – 40.9% (25.4%), Karaganda – 34.3% (46.5%), Almaty – 33.8% (35.7 %), Zhetysu - 31.4% (34.1%) regions and Shymkent city - 37.1% (42.9%). The lowest neglect is in the Atyrau region - 12.0% (17.5%).

In the country as a whole in 2022, the five-year survival rate of patients with CRC registered in 2018 decreased to 40.4% (2021 - 52.9% for those registered in 2017); there is a significant dispersion of indicators by region, from maximum – 56.1% (47.5%) in the Kyzylorda region, to minimum – 24.3% (51.5%) in the Aktobe region [5].

Screening of CRC screening is the systematic use of screening studies in an asymptomatic population. The purpose of screening is to identify people with abnormalities suggestive of CRC. These persons in the future need additional examination to clarify the diagnosis. Opportunistic screening is the non-systematic use of screening tests in routine medical practice. A screening program is much more challenging than an early detection program. At the same time, the success of the screening program is largely determined by the awareness of the population and medical workers about the possibilities of early diagnosis of CRC. The feasibility of a screening program is

determined by several factors that relate to the disease being screened, the screening test, the characteristics of the population, and the characteristics of the healthcare system.

The first factor is that the disease must be well understood, common enough in the target population to justify screening, have a recognizable early stage; treatment of the disease at an early stage should be more effective than at a later stage.

The second is that the test should be characterized by sufficient sensitivity, i.e. the ability to detect cancer among people with the disease; sufficient specificity - the probability that among people who do not have a disease, the test result will be negative; have a high positive predictive value (positive predictive value) or, in other words, the likelihood that people with a positive test result have the disease; have a high predictive value of a negative result (negative predictive value), i.e. the likelihood that people with a negative test result do not have the disease; security; low cost; and acceptability - the likelihood that people for whom this test is intended will agree to the examination (which to some extent depends on the awareness of the population about the possibilities and importance of early diagnosis).

The third factor is that the healthcare system should be ready for maximum screening test coverage of the target group, have the resources to confirm the diagnosis, appropriate treatment and follow-up of people with positive test results, and regularly conduct screening tests at regular intervals. At the same time, the benefits of screening must outweigh the potential physical and psychological harm and justify the financial costs of its implementation [14].

The factors most significant for the development of CRC are:

- the presence of chronic inflammatory bowel diseases, adenomatous polyps, cancer of other localization, etc.;

- family history (presence of one or two first-degree relatives with CRC or familial diffuse intestinal polyposis);

- the age of men and women over 50 years old, taking into account the fact that more than 90% of patients with colorectal cancer are people of this age (medium risk).

Age, regardless of gender, is an important risk factor for CRC. After the age of 50, the incidence of CRC increases from 8 to 160 per 100,000 population. Thus, people who have reached the age of 50, even in the absence of symptoms, constitute a moderate risk group for CRC.

The second category of increased risk of CRC (20%) is made up of persons with a genetic and family predisposition, suffering from chronic inflammatory bowel diseases, diffuse familial polyposis.

The high-risk CRC group is determined by the so-called Amsterdam criteria (the presence of malignant tumors in two generations, the presence of cancer in a first-line relative under the age of 50 years), in this case, CRC screening should be carried out after the age of 30 years [15].

The degree of individual risk of developing CRC is determined before screening to select the scope of studies and the frequency of their conduct.

The interval for oncological colorectal screening is 1 time in 2 years, target group: men and women aged 50-70 years, with the exception of persons registered at the dispensary for CRC and colon polyposis. At the same time, when forming the target group, one should take into account the absence of severe concomitant diseases, such as the presence of a common malignant neoplasm, cerebrovascular diseases in the stage of decompensation, chronic obstructive pulmonary disease with respiratory failure, cirrhosis of the liver, myocardial infarction with congestive heart failure, diabetes mellitus with vascular complications. and others, which are highly likely to lead to death in the next 10 years.

The first step in screening for CRC is the fecal occult blood test (FOBT). Traditionally, such methods include a benzidine test for occult blood in the feces. This is a biochemical method based on the assessment of pseudoperoxidase activity of hemoglobin. There is ample evidence that invitation to guaiac FOBT screening (gFOBT) reduces CRC mortality by approximately 15% in age-

matched average-risk populations.

To ensure the effectiveness of screening with gFOBT, the interval for screening under the national screening program should not exceed two years. To date, there is an immunochemical FOBT method - iFOBT, which is superior in efficiency to gFOBT in terms of the probability of detecting adenoma and cancer. iFOBT has improved analysis performance compared to gFOBT.

Immunochemical (immunochromatographic) examination of feces for occult blood - iFOBT or hemocult test is carried out for all men and women of the target group using an express method, which allows you to get a result within 3-5 minutes, without the participation of a medical worker. However, the evaluation of the test is carried out only by a medical worker in the PHC preventive department.

With a positive analysis of feces for occult blood, the second stage of colorectal screening is performed, which consists in endoscopic examination of the colon - total colonoscopy [7]. At the same time, in this case, this medical manipulation is of a therapeutic and diagnostic nature, since it allows one-stage removal of adenomatous polyps, which, according to various authors, occur in every third subject after 50 years of age. At the same time, women have 20% fewer polyps than men, but they have more right-sided lesions, which are more difficult to detect using fecal blood tests, because they are less traumatic [16].

What results were obtained from screening for CRC? In 2022, 937,859 men and women of the target group aged 50 to 70 years were examined during colorectal screening (in 2021 - 920,640) [5].

Colorectal screening revealed 325 cases of colorectal cancer in the reporting year, which is 114 cases more than in the previous year (211 cases). The detection rate increased from 0.23 to 0.35 per 1000 patients examined. Low detection of colorectal cancer was noted in Zhambyl, Karaganda, Kostanay, Kyzylorda, Mangistau, Turkestan - the worst result, East Kazakhstan regions, Astana city - from 0.07 to 0.30 per 1000 examined. The best result is in the North Kazakhstan region – 0.81 per 1000 examined. Compared to 2021, there was a decrease in the detection of colorectal cancer per 1000 people examined during screening in Karaganda (from 0.22 to 0.21), Kostanay (from 0.29 to 0.28), Mangistau (from 0.20 to 0.12) regions and Astana city (from 0.20 to 0.19).

Colon precancer (adenoma detection rate) was detected in 27.5% of patients who underwent colonoscopy (2021 – 22.8%). The detection rate of precancer in Akmola, Aktobe, Almaty (8.5% is the worst result), West Kazakhstan, Zhambyl, Kostanay, Kyzylorda, Mangistau, Pavlodar, North Kazakhstan, Turkestan regions and cities is lower than the national average Astana and Shymkent. The best result is 36.2% in Almaty city. It should be noted that the planned indicator for the detection of precancer of the colon and rectum in the country for 2022, according to the Comprehensive Plan, was 23.0% and was achieved.

In 2022, the proportion of patients identified during screening studies with early stages of malignant neoplasms (stages 0-I) was 26.2% during colorectal screening (in 2021 - 27.5%).

High early detection of colorectal cancer (above 30%) was noted in Akmola, West Kazakhstan, Karaganda, Kostanay, Kyzylorda, Turkestan regions and Astana city (57.1% - the best result). Not a single case of early cancer has been identified in the Mangistau region. Cases of cancer in stages III-IV detected during screening were registered in Akmola, Aktobe, Almaty, West Kazakhstan, Zhambyl, Karaganda, Kostanay, Mangistau regions and Almaty city. A total of 21 cases of colorectal cancer in stage III and 3 in stage IV were identified (in 2021 - 18 and 5, respectively) [5].

The complex analysis carried out allows us to conclude that satisfactory results of cancer screening can be achieved only with its proper organization, high quality of implementation, active participation in population screening, the use of highly sensitive tests and instrumental methods of preventive examination, as well as subsequent accurate diagnosis of identified tumors and

timely treatment. High-quality screening leads to early diagnosis of pedological diseases and malignant pathology in the early stages, which, in turn, increases the effectiveness of treatment and improves the prognosis of the disease. Target groups that, for one reason or another, do not participate in screening should be informed that there are no other methods other than screening that would reduce mortality from malignant neoplasms. Incidence and mortality rates from CC, BC and CRC clearly reflect the epidemiological situation with this pathology in the regions of our country.

LITERATURE

1 Dunlop KLA, Singh N, Robbins HA, Zahed H, Johansson M, Rankin NM, Cust AE. Implementation considerations for risk-tailored cancer screening in the population: A scoping review. Prev Med. 2024 Apr;181:107897. doi: 10.1016/j.ypmed.2024.107897.

2 Owens-Jasey C, Chen J, Xu R, Angier H, Huebschmann AG, Ito Fukunaga M, Chaiyachati KH, Rendle KA, Robien K, DiMartino L, Amante DJ, Faro JM, Kepper MM, Ramsey AT, Bressman E, Gold R. Implementation of Health IT for Cancer Screening in US Primary Care: Scoping Review. JMIR Cancer. 2024 Apr 30;10:e49002. doi: 10.2196/49002.

3 Ebbert JO, Hawk ET, Chambers CV, Tempero MA, Fishman EK, Ravenell JE, Beer TM, Rego SP. Multi-cancer early detection tests: Attributes for clinical implementation. Cancer Biomark. 2025 Feb;42(2):18758592241297849. doi: 10.1177/18758592241297849.

4 Ilbawi A., Varghese Ch., Loring B., Ginsburg O., Corbex M. Under the overall direction of Krug E. and Varghese Ch. Guide to Cancer Early Diagnosis. World Health Organization, 2017; 48 p.

5 Kaidarova D.R., Shatkovskaya O.V., Ongarbayev B.T., Seisenbayeva G.T., Azhmagambetova A.E., Zhylkaidarova A.Zh., Lavrentieva I.K., Sagi M.S. Indicators of the oncology service of the Republic of Kazakhstan, 2022: statistical and analytical materials. – Almaty, 2023. – 430 p.

6 Fontham ETH, Wolf AMD, Church TR, Etzioni R, Flowers CR, Herzig A, Guerra CE, Oeffinger KC, Shih YT, Walter LC, Kim JJ, Andrews KS, DeSantis CE, Fedewa SA, Manassaram-Baptiste D, Saslow D, Wender RC, Smith RA. Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society. CA Cancer J Clin. 2020 Sep;70(5):321-346. doi: 10.3322/caac.21628.

7 Order of the Acting Minister of Healthcare of the Republic of Kazakhstan No. KR DSM-174/2020 dated October 30, 2020, "On Approval of Target Groups of Persons Subject to Screening Examinations, as well as the Rules, Scope, and Frequency of Conducting Such Examinations."

8 Order of the Minister of Healthcare of the Republic of Kazakhstan No. 62 dated August 16, 2024, "On Amending the Order of the Acting Minister of Healthcare of the Republic of Kazakhstan No. KR DSM-174/2020 dated October 30, 2020, 'On Approval of Target Groups of Persons Subject to Screening Examinations, as well as the Rules, Scope, and Frequency of Conducting Such Examinations'"

9 Order of the Acting Minister of Healthcare of the Republic of Kazakhstan No. 71 dated July 29, 2025, "On Amending the Order of the Acting Minister of Healthcare of the Republic of Kazakhstan No. KR DSM-174/2020 dated October 30, 2020, 'On Approval of Target Groups of Persons Subject to Screening Examinations, as well as the Rules, Scope, and Frequency of Conducting Such Examinations'"

10 Perkins RB, Guido RL, Saraiya M, Sawaya GF, Wentzensen N, Schiffman M, Feldman S. Summary of Current Guidelines for Cervical Cancer Screening and Management of Abnormal Test Results: 2016-2020. J Womens Health (Larchmt). 2021 Jan;30(1):5-13. doi: 10.1089/jwh.2020.8918.

11 Abdolell M, Payne JI, Caines J, Tsuruda K, Barnes PJ, Talbot PJ, Tong O, Brown P, Rivers-

Bowerman M, Iles S. Assessing breast cancer risk within the general screening population: developing a breast cancer risk model to identify higher risk women at mammographic screening. Eur Radiol. 2020 Oct;30(10):5417-5426. doi: 10.1007/s00330-020-06901-x.

12 Melnik I, Rapson Y, Gropstein A, Sharon E. Different approaches to mammography as a screening tool for breast cancer. Harefuah. 2022 Feb;161(2):121-124.

13 Mann R.M., Athanasiou A., Baltzer P.A.T. et al. Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). Eur Radiol. 2022 Jun;32(6):4036-4045. doi: 10.1007/s00330-022-08617-6.

14 Kashin S.V., Nehajkova N.V., Zav'jalov D.V. i dr. Skrining kolorektal'nogo raka: obshhaja situacija v mire i rekomendovannye standarty kachestva kolonoskopii. Dokazatel'naja gastrojenterologija. 2017;6(4):32-52 (In Russ.).

15 Samadder N.J., Smith K.R., Wong J. et al. Cancer risk in families fulfilling the Amsterdam Criteria for Lynch syndrome. JAMA Oncol. 2017 Dec 1;3(12):1697-1701. doi: 10.1001/jamaoncol.2017.0769.

16 Hultcrantz R. Aspects of colorectal cancer screening, methods, age and gender. J Intern Med. 2021 Apr;289(4):493-507. doi: 10.1111/joim.13171.

THE SCIENTIFIC DISCOURSE ON THE CONTEMPORARY PHARMACOTHERAPEUTIC ARMAMENTARIUM OF DUCHENNE MUSCULAR DYSTROPHY: AN INTEGRATIVE ANALYSIS OF THE PHARMACOLOGICAL CLASSIFICATION OF THERAPEUTIC AGENTS, MOLECULAR PHARMACOLOGY, PHARMACODYNAMIC AND PHARMACOKINETIC DETERMINANTS, ADVERSE DRUG REACTIONS, TOXICOLOGICAL PROFILES AND COMPREHENSIVE PHARMACOSAFETY CONSIDERATIONS

Nodar Sulashvili
MD, PhD, Doctor of Pharmaceutical and Pharmacological Sciences In Medicine, Professor of Medical and Clinical Pharmacology of International School of Medicine at Alte University; Associate Professor of Biochemistry and Pharmacology Direction of School of Health Sciences at The University of Georgia; Invited Professor/Lecturer of Scientific Research-Skills Center at Tbilisi State Medical University; Invited Lecturer of Clinical Pharmacology Department at Tbilisi State Medical University; Scientific Researcher, Best, Top Academic Scientist: Georgia, Europe, Worldwide; Tbilisi, Georgia; Scientific Researcher of Yerevan State Medical University After Mkhitar Heratsi, Yerevan, Republic of Armenia; Former Clinical Pharmacologist of London Lewisham University Hospital, London, England, United Kingdom of Great Britain and Northern Ireland; https://orcid.org/0000-0002-9005-8577
Margarita Beglaryan
MD, PhD, Doctor of Pharmaceutical Sciences, Academician, Professor of Yerevan State Medical University After Mkhitar Heratsi, Head of the Department of Pharmaceutical

Management of Yerevan State Medical University After Mkhitar Heratsi; Yerevan, Republic of Armenia; https://orcid.org/0000-0003-3697-6390

Diego Rada Fernandez de Jauregui

PhD and BSc in Pharmacy, MSc in Research Methodology in Health Sciences, Professor, Pharmacy Faculty, University of the Basque Country, Spain. https://orcid.org/0000-0002-9109-1088

Irine Pkhakadze

MD, PhD, Doctor of Medical Sciences, Professor of Faculty of Medicine at Akaki Tsereteli State University, Kutaisi, Georgia;

Giorgi Gabisonia

MD, MPH, PhDc, Dean of the International School of Medicine at Alte University, Tbilisi, Georgia.

Marika Sulashvili

MD, Doctor of Family Medicine, Invited Lecturer of Department of Molecular and Medical Genetics at Tbilisi State Medical University, Invited Professor of Biochemistry and Molecular and Medical Genetics at The University of Georgia. Tbilisi, Georgia. https://orcid.org/0000-0002-6338-4262.

Tamar Kraveishvili

MD, Head of The Preclinical Department of International School of Medicine at Alte University; Public Health Institute of Georgia (PHIG); Tbilisi, Georgia https://orcid.org/0009-0009-2473-2568

Nana Gorgaslidze

MD, PhD, Doctor of Pharmaceutical Sciences, Academician, Professor of Tbilisi State Medical University, Head of The Department of Social and Clinical Pharmacy, Tbilisi, Georgia. https://orcid.org/0000-0002-4563-5224

Nato Alavidze

MD, PhD, Doctor of Pharmaceutical Sciences, Professor of Akaki Tsereteli State University, Faculty of Medicine, Department of Stomatology and Pharmacy, Kutaisi, Georgia; https://orcid.org/0000-0001-6695-5924

Naira Chichoyan

MD, PhD, Doctor of Pharmaceutical Sciences, Professor of Yerevan State Medical University After Mk. Heratsi; Head of the Department of Pharmacognosy at Yerevan State Medical University After Mkhitar Heratsi; Yerevan, Republic of Armenia. https://orcid.org/0000-0003-3905-3372

Marina Giorgobiani

MD, PhD, Doctor of Medical Sciences, Academician, Professor of Tbilisi State Medical University, Faculty of Public Health; Department of Hygiene and Medical Ecology, Tbilisi, Georgia. https://orcid.org/0000-0003-0686-5227

Ada (Adel) Tadevosyan

MD, PhD, Doctor of Medical Sciences, Academician, Professor of Yerevan State Medical University After Mkhitar Heratsi; Academician and Full Member of the International Academy of Sciences of Ecology and Life Safety (MANEB), Academician of the Republic Armenia Law Academy, Member of the Association of Psychiatrists of Armenia, World Association of Psychiatrists, International Association for Traumatic Stress, World Association for Biological Psychiatry, International Association "Stress and Behavior",

Licensed Psychiatrist, Psychotherapist, Public Health Organizer; Tbilisi-Georgia, Yerevan-Armenia, Los Angeles-USA;

Igor Seniuk

PhD, Doctor of Pharmaceutical Sciences, Dean of faculty of Pharmacy at National University of Pharmacy of Ukraine, Associate Professor of Biological Chemistry Department at National University of Pharmacy, Kharkiv, Ukraine. https://orcid.org/0000-0003-3819-7331

Vira Kravchenko

MD, PhD, Doctor of Biological Sciences, Academician, Professor, Head of The Biological Chemistry Department at National University of Pharmacy of Ukraine, Kharkiv, Ukraine. https://orcid.org/0000-0001-6335-2490

Tamar Okropiridze

MD, DDM, PhD, Doctor of Medical Sciences, Academician, Professor of the Division of Dentistry of International School of Medicine at Alte University; Invited Professor of Dentistry Department of The School of Health Sciences at The University of Georgia, Tbilisi, Georgia.

Olga Shapoval

MD, PhD, Doctor of Pharmaceutical Sciences, Associate Professor of Department of Pharmacology and Medical Prescription at Kharkiv State Medical University, Kharkiv, Ukraine; https://orcid.org/0000-0002-1971-2863

Mohini Yadav

International School of Medicine at Alte University; Tbilisi, Georgia;

Udgam Yaduvanshi

International School of Medicine at Alte University; Tbilisi, Georgia;

Lali Patsia

MD, PhD, Doctor of Medical Sciences, Doctor Cardiologist at Republican Hospital, Professor of International School of Medicine at Alte University; Invited Professor of Tbilisi State Medical University, Tbilisi, Georgia.

Kakhaber Robakidze

MD, PhD, Doctor of Medical Sciences, Academician, Professor of International School of Medicine at Alte University; Professor of Faculty of Medicine at Caucasus International University; Doctor-Professor of National Health Center Named After Academician O. Gudushauri; Tbilisi, Georgia;

David Aphkhazava

PhD, Doctor of Biological Sciences, Professor of Biochemistry of International School of Medicine at Alte University; Professor of Biochemistry at The University of Georgia, Tbilisi, Georgia. Orcid: https://orcid.org/0000-0001-6216-6477

Corresponding Author: Dr. Professor Nodar Sulashvili

ABSTRACT

Duchenne muscular dystrophy (DMD) constitutes a paradigmatic X-linked neuromuscular disorder characterized by pathogenic alterations in the *DMD* gene, disruption of dystrophin expression, progressive sarcolemmal instability, chronic myofiber degeneration, persistent inflammatory remodeling, impaired regenerative capacity, and inexorable deterioration of skeletal, respiratory, and cardiac muscle function. The contemporary pharmacotherapeutic armamentarium of DMD

has undergone substantial expansion from predominantly glucocorticoid-based management toward an increasingly mechanistically stratified therapeutic paradigm encompassing corticosteroid pharmacotherapy, exon-skipping antisense oligonucleotides, histone deacetylase inhibition, micro-dystrophin gene replacement, and multidimensional pharmacological management of disease-associated cardiopulmonary, skeletal, endocrine, and functional complications. This scientific discourse provides an integrative pharmacological appraisal of currently established and emerging therapeutic modalities, emphasizing therapeutic classification, molecular mechanisms of action, pharmacodynamic determinants, pharmacokinetic characteristics, genotype-dependent treatment selection, adverse drug reactions, toxicological liabilities, pharmacological interactions, and comprehensive pharmacosafety considerations. Glucocorticoids remain fundamental components of DMD pharmacotherapy because of their clinically meaningful effects on muscle strength, motor performance, ambulation, and pulmonary function. Prednisone/prednisolone and deflazacort exert pleiotropic genomic and non-genomic corticosteroid effects through glucocorticoid-receptor signaling, modulating transcriptional networks involved in inflammation, immune activation, muscle homeostasis, and tissue remodeling. Vamorolone represents a pharmacologically distinctive corticosteroid derivative designed to retain glucocorticoid-mediated therapeutic activity while modifying receptor-dependent transactivation and transrepression profiles, thereby potentially altering the conventional corticosteroid toxicity spectrum. Nevertheless, chronic corticosteroid exposure necessitates systematic surveillance for growth impairment, weight gain, behavioral and metabolic abnormalities, hypertension, cataracts, osteoporosis, infection susceptibility, and hypothalamic–pituitary–adrenal axis suppression, rendering individualized dose optimization and adrenal-stress management integral components of pharmacosafety. Molecularly targeted exon-skipping therapies represent a genotype-dependent pharmacological strategy designed to manipulate pre-mRNA splicing and restore the translational reading frame of selected *DMD* transcripts. Eteplirsen, golodirsen, viltolarsen, and casimersen are phosphorodiamidate morpholino oligomer or related antisense oligonucleotide therapeutics directed against specific exons, respectively enabling exon 51, exon 53, or exon 45 skipping in genetically eligible patients. Their pharmacodynamic objective is the induction of dystrophin production through sequence-specific hybridization with dystrophin pre-mRNA. Their pharmacokinetic behavior is characterized by distribution to tissues, intracellular uptake, metabolic stability relative to conventional small molecules, and predominantly non-CYP-mediated elimination pathways. Pharmacovigilance nevertheless remains essential, particularly regarding renal safety, infusion-related reactions, hypersensitivity, hepatic or laboratory abnormalities, and the uncertainty associated with surrogate biomarker-mediated therapeutic assessment. Gene-replacement therapy has introduced a fundamentally different pharmacological modality through systemic delivery of a recombinant adeno-associated virus vector encoding a micro-dystrophin transgene. Delandistrogene moxeparvovec-rokl is intended to facilitate expression of a functional micro-dystrophin protein within skeletal and cardiac muscle, thereby addressing the underlying protein-deficiency phenotype rather than merely modifying downstream pathophysiology. Its pharmacological complexity encompasses vector biodistribution, cellular transduction, transgene expression, pre-existing anti-vector immunity, complement activation, hepatotoxicity, thrombocytopenia, and potentially severe immune-mediated adverse events. Consequently, patient selection, baseline serological assessment, corticosteroid immunomodulation, hepatic monitoring, and long-term pharmacovigilance constitute essential elements of responsible gene-therapy deployment. Givinostat, an orally administered histone deacetylase inhibitor, further expands the mechanistic spectrum of DMD pharmacotherapy by modulating epigenetic and transcriptional processes implicated in muscle pathology, inflammation, fibrosis, and regenerative dysfunction. Its pharmacodynamic effects are therefore conceptually distinct from corticosteroid

receptor signaling, exon-specific RNA modulation, and gene replacement. Monitoring of hematological parameters, lipid metabolism, hepatic function, gastrointestinal tolerability, and other clinically relevant laboratory variables is necessary to mitigate treatment-associated toxicity. Beyond disease-modifying therapy, comprehensive DMD pharmacotherapy requires coordinated management of cardiomyopathy, respiratory insufficiency, osteoporosis, endocrine consequences, pain, contractures, and other secondary complications. Renin–angiotensin-system inhibitors, mineralocorticoid-receptor antagonists, β-adrenergic blockers, diuretics, and other cardiovascular agents may be incorporated according to individualized cardiac phenotype and guideline-directed indications, while bone-directed interventions and nutritional/endocrine therapies address corticosteroid-associated and disease-related skeletal vulnerability. Respiratory pharmacotherapy remains predominantly complication-oriented rather than disease-modifying. The contemporary DMD pharmacotherapy represents an increasingly sophisticated convergence of molecular genetics, RNA therapeutics, epigenetic pharmacology, gene therapy, immunopharmacology, and conventional clinical pharmacology. Optimal therapeutic decision-making requires integration of genotype, age, ambulatory status, disease stage, organ-system involvement, pharmacokinetic variability, treatment response, and cumulative toxicity. The emerging paradigm consequently extends beyond efficacy-centered prescribing toward precision pharmacotherapy, longitudinal pharmacovigilance, individualized benefit–risk assessment, and proactive pharmacosafety optimization throughout the lifelong clinical trajectory of DMD.

Keywords: Duchenne muscular dystrophy; Pharmacotherapy; Molecular pharmacology; Pharmacodynamics; Pharmacokinetics; Pharmacovigilance; Pharmacosafety.

INTRODUCTION

Duchenne muscular dystrophy (DMD) represents one of the most consequential inherited disorders within the spectrum of muscular dystrophies and constitutes a prototypical model of progressive genetic muscle disease in which molecular pathology, cellular degeneration, tissue remodeling, systemic complications, and therapeutic pharmacology are intrinsically interconnected. Characterized principally by pathogenic variants in the *DMD* gene and consequent deficiency or functional absence of dystrophin, DMD produces a continuously evolving phenotype involving skeletal muscle, respiratory musculature, myocardium, and multiple secondary physiological systems. The disorder is therefore no longer appropriately conceptualized solely as a primary skeletal-muscle disease. Rather, it represents a multisystem pathobiological entity in which defective cytoskeletal–sarcolemmal integrity initiates a cascade of mechanical, biochemical, inflammatory, metabolic, fibrotic, and electrophysiological abnormalities that progressively compromise functional capacity and survival. The therapeutic implications of this complex biology have been profound, driving an evolution from predominantly symptomatic management toward molecularly informed pharmacotherapy directed at distinct levels of disease pathogenesis.

The fundamental molecular abnormality underlying DMD is disruption of the dystrophin-associated structural and signaling network. Dystrophin is a large cytoskeletal protein that links intracellular actin filaments to the dystrophin-associated glycoprotein complex situated within the sarcolemma. Through this molecular architecture, dystrophin contributes to stabilization of the muscle-cell membrane during repetitive contraction and helps coordinate mechanical force transmission between the intracellular cytoskeleton and extracellular matrix. Loss or profound reduction of functional dystrophin renders myofibers unusually susceptible to contraction-induced mechanical stress. Recurrent sarcolemmal disruption facilitates abnormal calcium influx, activation of proteolytic systems, mitochondrial dysfunction, oxidative stress, necrotic myofiber injury, and release of intracellular components that amplify inflammatory signaling. Progressive cycles of degeneration and incomplete regeneration ultimately result in replacement of

contractile tissue by adipose and fibrotic tissue. The pharmacological consequences of this sequence are substantial because therapeutic intervention may theoretically occur at multiple levels, including restoration of dystrophin expression, modification of RNA processing, suppression of pathological inflammation, modulation of epigenetic transcriptional pathways, correction of secondary cardiopulmonary dysfunction, preservation of skeletal integrity, and attenuation of treatment-related toxicity.

The *DMD* gene is among the largest genes in the human genome, and its molecular complexity contributes substantially to the genetic heterogeneity of DMD. Pathogenic variants include deletions, duplications, and smaller sequence alterations that can disrupt the translational reading frame or otherwise impair dystrophin production. This genotype–phenotype relationship has become increasingly important in modern therapeutic pharmacology because several contemporary interventions are explicitly dependent upon the precise molecular defect. Consequently, the pharmacotherapy of DMD cannot be regarded as a uniform therapeutic algorithm applicable to all patients. Instead, treatment selection increasingly requires molecular characterization of the pathogenic variant, determination of exon architecture, assessment of reading-frame consequences, evaluation of age and disease stage, determination of ambulatory status, characterization of cardiac and respiratory involvement, and consideration of previous or concomitant therapies. Precision pharmacotherapy in DMD therefore represents the convergence of molecular diagnostics and pharmacological decision-making.

Historically, the pharmacological management of DMD was dominated by glucocorticoids. Prednisone or prednisolone and deflazacort established the therapeutic foundation for disease modification through their capacity to preserve muscle strength and prolong functional performance. Although the precise molecular basis for the beneficial effects of glucocorticoids in DMD remains incompletely resolved, their pharmacology extends beyond simple suppression of generalized inflammation. Activation of intracellular glucocorticoid receptors modifies transcriptional networks involved in immune regulation, inflammatory signaling, cellular metabolism, muscle homeostasis, and tissue remodeling. Glucocorticoid treatment can alter the balance between inflammatory injury and muscle repair, potentially attenuating secondary mechanisms that accelerate myofiber degeneration. Nevertheless, the pharmacological efficacy of chronic corticosteroid therapy is accompanied by substantial systemic toxicity, including growth suppression, weight gain, behavioral abnormalities, hypertension, glucose dysregulation, osteoporosis, vertebral fractures, cataract formation, immunosuppression, and hypothalamic–pituitary–adrenal axis suppression. The therapeutic history of DMD consequently illustrates a fundamental principle of clinical pharmacology: the clinical value of a pharmacological intervention is determined not only by its biological efficacy but also by the magnitude, reversibility, detectability, and manageability of its toxicity.

Deflazacort subsequently provided an alternative corticosteroid strategy characterized by a pharmacological profile that differs from conventional prednisone therapy. Its introduction into DMD treatment illustrates the importance of structural modification within a therapeutic class to influence pharmacodynamic and adverse-effect profiles. The clinical utility of deflazacort is associated with preservation of motor function and potential differences in weight-related and metabolic toxicity compared with prednisone, although its own adverse-effect spectrum includes clinically important risks, particularly cataracts, growth effects, adrenal suppression, and skeletal complications. Accordingly, corticosteroid selection in DMD is not simply a matter of identifying the most potent anti-inflammatory agent; rather, it involves individualized assessment of therapeutic response, anthropometric trajectory, behavioral effects, bone health, ophthalmological status, endocrine function, and long-term treatment adherence.

The development of vamorolone represents a further conceptual evolution in corticosteroid pharmacology. Rather than simply introducing another conventional glucocorticoid, vamorolone

was designed to modify glucocorticoid-receptor pharmacology with the objective of retaining therapeutically desirable activity while potentially reducing selected corticosteroid-associated toxicities. Its development reflects an increasingly sophisticated approach to ligand–receptor pharmacology in which therapeutic efficacy and adverse-effect profiles are conceptualized as potentially separable through molecular modulation of receptor signaling. Vamorolone therefore occupies an important position within the contemporary DMD pharmacotherapeutic armamentarium, particularly because its pharmacological rationale illustrates how medicinal chemistry can be used to address the limitations of long-established therapeutic classes.

The emergence of antisense oligonucleotide therapy fundamentally transformed the conceptual framework of DMD pharmacotherapy. Conventional small-molecule drugs generally modify protein function, receptor signaling, enzymatic activity, or downstream physiological processes. Exon-skipping therapies operate at the level of RNA processing and therefore intervene much closer to the molecular origin of dystrophin deficiency. Eteplirsen, golodirsen, viltolarsen, and casimersen exemplify sequence-specific antisense pharmacology directed against selected dystrophin exons. By hybridizing with complementary sequences within dystrophin pre-mRNA, these agents can alter spliceosome recognition and promote omission of a specific exon during mRNA processing. The resulting transcript may restore the translational reading frame in genetically eligible patients and permit synthesis of a truncated but potentially functional dystrophin protein.

The pharmacological significance of exon-skipping therapy extends beyond its molecular mechanism. These agents represent genotype-dependent therapeutics in which the genetic substrate determines whether a particular drug can theoretically exert its intended molecular effect. Eteplirsen is directed toward exon 51 skipping, whereas golodirsen and viltolarsen target exon 53, and casimersen targets exon 45. Consequently, genetic testing is not merely diagnostic but becomes an integral component of pharmacotherapy. The therapeutic paradigm thereby shifts from phenotype-centered prescribing toward genotype-informed intervention. This transformation has broad implications for clinical pharmacology, because drug eligibility, anticipated pharmacodynamic response, biomarker interpretation, and treatment continuation may all depend upon molecular characteristics that traditionally belonged to the domain of genetics rather than pharmacology.

Antisense oligonucleotides also present distinctive pharmacokinetic and pharmacosafety challenges. Unlike many orally administered small molecules, these macromolecular therapeutics exhibit physicochemical properties that strongly influence tissue distribution, cellular uptake, metabolism, and elimination. Their therapeutic activity depends on achieving adequate intracellular exposure within relevant muscle tissues and establishing sufficient interaction with the target RNA. Pharmacokinetic behavior may consequently be influenced by molecular architecture, tissue binding, renal handling, intracellular trafficking, and repeated intravenous administration. Moreover, pharmacodynamic exposure cannot necessarily be inferred from conventional plasma concentration measurements alone. The relationship between systemic exposure, tissue penetration, exon-skipping activity, dystrophin expression, and clinical phenotype is inherently complex. These considerations emphasize the importance of integrating pharmacokinetic, pharmacodynamic, molecular, and clinical endpoints when evaluating RNA-targeted therapies.

The expansion of molecular therapeutics was followed by the development of gene-replacement therapy, representing another fundamental shift in therapeutic philosophy. Delandistrogene moxeparvovec-rokl employs an adeno-associated viral vector platform to deliver a transgene encoding a micro-dystrophin protein to muscle tissue. Whereas corticosteroids primarily modify downstream disease processes and exon-skipping therapies manipulate dystrophin RNA processing, gene replacement seeks to introduce a genetic template capable of promoting

expression of a dystrophin-derived therapeutic protein. This strategy therefore operates at a distinct biological level and introduces pharmacological characteristics more commonly associated with gene therapy than conventional medicinal products.

Gene-replacement pharmacology is intrinsically multidimensional. Following systemic administration, vector particles undergo biodistribution and interact with target tissues, followed by cellular entry, intracellular trafficking, nuclear delivery, and transgene expression. Therapeutic performance is influenced by vector biology, tissue tropism, host immunity, pre-existing anti-vector antibodies, complement activation, corticosteroid-mediated immunomodulation, hepatic exposure, and the persistence of transgene expression. These factors generate a therapeutic profile in which conventional concepts of dose, concentration, clearance, and reversibility require reinterpretation. Unlike a conventional small-molecule drug that can usually be rapidly discontinued when toxicity emerges, gene therapy may produce prolonged biological effects after a single administration. Consequently, pharmacovigilance must extend beyond the conventional treatment period and incorporate longitudinal monitoring for immune-mediated, hepatic, hematological, and other potentially delayed adverse outcomes.

Givinostat provides another mechanistically differentiated component of modern DMD therapy. As a histone deacetylase inhibitor, givinostat acts upon epigenetic regulatory mechanisms rather than directly replacing dystrophin or altering a specific exon sequence. Histone acetylation and deacetylation participate in regulation of chromatin accessibility and transcriptional activity, and dysregulated epigenetic signaling may contribute to inflammatory, fibrotic, and regenerative abnormalities within dystrophic muscle. Pharmacological inhibition of histone deacetylase activity can therefore influence transcriptional programs relevant to muscle pathology. The therapeutic development of givinostat demonstrates the increasing recognition that DMD pathology is not exclusively determined by the absence of dystrophin itself but is also shaped by downstream molecular networks that can potentially be pharmacologically modulated.

The emergence of these mechanistically heterogeneous therapies necessitates a more sophisticated pharmacological taxonomy of DMD treatment. Contemporary classification should distinguish corticosteroid disease-modifying therapy, RNA-targeted exon-skipping agents, epigenetic modulators, gene-replacement therapy, and pharmacotherapy directed toward secondary systemic complications. Such classification is clinically meaningful because these therapeutic groups differ substantially in mechanism of action, route of administration, pharmacokinetic behavior, pharmacodynamic endpoints, treatment eligibility, adverse-effect profiles, monitoring requirements, and evidence standards. A comprehensive pharmacological analysis must therefore move beyond a simple enumeration of available medications and instead examine the mechanistic continuum extending from molecular target engagement to clinically meaningful functional outcomes.

The pathophysiology of DMD further demonstrates why disease-modifying treatment cannot be separated from supportive pharmacotherapy. Progressive myocardial involvement represents one of the major determinants of morbidity and mortality. Dystrophin deficiency within cardiomyocytes promotes membrane instability, cellular injury, myocardial fibrosis, electrical abnormalities, and progressive dilated cardiomyopathic remodeling. Consequently, cardiovascular pharmacotherapy constitutes an essential component of comprehensive DMD management. Agents affecting the renin–angiotensin–aldosterone system, β-adrenergic signaling, mineralocorticoid-receptor activity, preload, afterload, and myocardial remodeling may be incorporated according to individual cardiac phenotype and contemporary heart-failure principles. ACE inhibitors, angiotensin-receptor blockers, mineralocorticoid-receptor antagonists, and β-blockers have consequently become important elements of cardioprotective pharmacology in DMD. Their use illustrates the distinction between DMD-specific pharmacotherapy and indication-based management of disease complications: these agents do not correct the primary

dystrophin defect but may substantially influence the clinical consequences of progressive myocardial disease.

Cardiac pharmacotherapy in DMD also requires rigorous consideration of pharmacokinetic and pharmacodynamic interactions. Renal function, serum potassium, blood pressure, cardiac rhythm, ventricular function, and concomitant medications can substantially modify the benefit–risk profile of renin–angiotensin-system inhibitors and mineralocorticoid-receptor antagonists. B-blockers require attention to heart rate, conduction abnormalities, blood pressure, fatigue, and respiratory status. Diuretics may be necessary in advanced congestion but introduce risks of electrolyte imbalance, volume depletion, and renal dysfunction. Therefore, pharmacosafety in DMD cannot be reduced to monitoring adverse events associated with disease-specific drugs; it must encompass the entire pharmacological ecosystem within which multiple therapeutic classes are administered over prolonged periods.

Respiratory deterioration represents another major component of DMD progression. Weakness of the diaphragm and accessory respiratory muscles, impaired cough effectiveness, sleep-related hypoventilation, and progressive reduction in ventilatory reserve create an evolving respiratory phenotype. Pharmacological intervention is generally less central than non-invasive ventilatory support and mechanical cough assistance, yet medications may be required for specific respiratory complications. Bronchodilators, mucolytic therapies, antimicrobial treatment, and other agents may be used according to clinical indication rather than as primary disease-modifying interventions. This distinction is critical because indiscriminate pharmacological treatment of respiratory symptoms may expose patients to unnecessary adverse effects without addressing the underlying neuromuscular ventilatory impairment.

Skeletal and endocrine complications further broaden the pharmacological requirements of DMD. Chronic glucocorticoid therapy, reduced weight-bearing activity, muscle weakness, and intrinsic abnormalities of bone metabolism contribute to low bone mineral density and increased risk of vertebral and long-bone fractures. Accordingly, vitamin D and calcium adequacy, endocrine surveillance, and appropriately selected bone-directed therapies constitute important elements of long-term pharmacological management. Intravenous bisphosphonate therapy has particular relevance in the management of clinically significant osteoporosis and vertebral fractures in affected patients. The pharmacological approach must nevertheless account for age, skeletal maturity, renal function, calcium and vitamin D status, fracture history, and the cumulative consequences of long-term corticosteroid exposure.

Pain, contractures, sleep disturbance, gastrointestinal symptoms, and other functional complications may additionally require pharmacological intervention. However, symptomatic treatment should remain integrated within a multidisciplinary strategy rather than substituting for disease-modifying therapy or rehabilitative care. Analgesics, antispasticity agents, and other symptom-directed medications may provide clinically meaningful benefits but can also generate sedation, gastrointestinal toxicity, hepatotoxicity, renal effects, or respiratory compromise. In patients with advanced neuromuscular weakness, even apparently conventional medications may acquire a different risk profile because respiratory reserve, mobility, nutritional status, and polypharmacy can alter pharmacological tolerance.

The concept of pharmacosafety is therefore particularly important in DMD. Patients frequently require lifelong exposure to multiple medications, beginning during childhood and continuing through adolescence and adulthood. This prolonged therapeutic trajectory creates cumulative risks that may not be apparent in short-term clinical trials. Corticosteroid-associated osteoporosis, metabolic abnormalities, adrenal suppression, cataracts, and growth effects can emerge progressively. RNA-targeted therapies require repeated administration and ongoing surveillance for organ-specific toxicity. Gene therapy introduces the possibility of prolonged or irreversible biological exposure and immune-mediated complications. Epigenetic therapy requires laboratory

monitoring for hematological, metabolic, hepatic, and other treatment-associated abnormalities. Cardiovascular medications require surveillance of renal function, electrolytes, blood pressure, and cardiac status. The pharmacosafety framework must therefore be longitudinal, proactive, and multidimensional.

Pharmacovigilance assumes particular importance because rare diseases frequently present evidentiary challenges that differ from those encountered in common disorders. Limited patient populations, heterogeneous genotypes, variable disease trajectories, small clinical trials, surrogate endpoints, evolving regulatory frameworks, and long-term uncertainty can complicate the estimation of uncommon or delayed adverse events. Real-world evidence, post-marketing surveillance, disease registries, longitudinal observational cohorts, and structured adverse-event reporting therefore become indispensable components of therapeutic evaluation. In DMD, pharmacovigilance must encompass not only the identification of adverse drug reactions but also assessment of treatment discontinuation, cumulative toxicity, drug–drug interactions, adherence, dose modifications, and clinically meaningful changes in functional outcomes.

The pharmacodynamic assessment of DMD therapies also presents distinctive challenges. Traditional pharmacology frequently links drug exposure to measurable receptor occupancy, enzyme inhibition, or physiological response. In DMD, however, pharmacodynamic biomarkers may operate across several biological scales. Molecular endpoints can include exon-skipping efficiency, dystrophin transcript production, dystrophin protein expression, or micro-dystrophin expression following gene therapy. Cellular and tissue-level endpoints may include inflammatory activity, fibrosis, muscle regeneration, and membrane integrity. Functional endpoints include timed motor performance, muscle strength, ambulation, respiratory function, and cardiac performance. The clinical interpretation of pharmacodynamic biomarkers consequently requires an understanding of their relationship to actual disease modification. An increase in a molecular biomarker does not automatically establish a proportional improvement in long-term clinical outcomes, and this distinction is central to rigorous therapeutic evaluation.

Pharmacokinetics is similarly complex in the context of DMD. Developmental changes during childhood and adolescence may modify drug distribution, metabolism, and clearance. Progressive loss of skeletal muscle mass can alter body composition and complicate weight-based dosing. Reduced mobility, nutritional abnormalities, changes in renal physiology, corticosteroid exposure, and concomitant medications may further modify pharmacokinetic behavior. For macromolecular therapies, tissue distribution and intracellular trafficking may be more clinically relevant than conventional plasma pharmacokinetic parameters. For gene therapy, vector biodistribution and transgene expression represent fundamentally different dimensions of therapeutic exposure. These considerations demonstrate that pharmacokinetic characterization must be interpreted within the biological context of progressive neuromuscular disease rather than as an isolated mathematical description of drug concentration over time.

Drug interactions constitute another important pharmacological consideration. Polypharmacy becomes increasingly prevalent as DMD progresses and cardiac, respiratory, skeletal, endocrine, and symptomatic complications accumulate. Pharmacodynamic interactions may occur when multiple agents affect blood pressure, cardiac conduction, renal function, electrolyte balance, central nervous system activity, or immune responses. Pharmacokinetic interactions may arise through alterations in metabolic enzymes, transporters, renal elimination, or gastrointestinal absorption. Corticosteroids additionally interact conceptually with gene therapy because immunomodulatory regimens may be deliberately incorporated into treatment protocols. Therefore, comprehensive medication reconciliation and individualized benefit–risk assessment are essential components of DMD pharmacotherapy.

The therapeutic evolution of DMD also raises important questions regarding treatment sequencing and combination therapy. The increasing availability of mechanistically distinct

disease-modifying interventions creates opportunities for complementary pharmacological strategies but simultaneously generates uncertainty regarding optimal combinations, treatment initiation thresholds, switching criteria, and long-term sequencing. Corticosteroids may remain foundational even when molecular therapies are introduced, while cardiac and respiratory pharmacotherapies may continue independently of genotype-specific disease-modifying interventions. The possibility of combining therapies acting at different molecular levels is scientifically compelling but requires rigorous evaluation of additive efficacy, overlapping toxicity, pharmacokinetic compatibility, immunological consequences, and long-term clinical benefit.

Another defining characteristic of contemporary DMD pharmacology is the increasing importance of personalized therapeutic decision-making. The traditional paradigm of "one disease, one treatment" is increasingly incompatible with a disorder characterized by substantial genetic and phenotypic heterogeneity. Two patients carrying different pathogenic variants may have entirely different eligibility for exon-skipping therapy. Two patients with the same genotype may differ substantially in age, ambulatory capacity, cardiac involvement, respiratory reserve, steroid tolerance, and comorbidity burden. Consequently, precision pharmacotherapy requires integration of molecular diagnosis with longitudinal phenotyping and individualized pharmacological surveillance.

The regulatory development of DMD therapies has also influenced the evidentiary architecture of modern pharmacology. Because DMD is a rare, progressive, and life-limiting disease, conventional large randomized clinical trials may be difficult to conduct, particularly for genetically restricted therapies. Regulatory pathways have consequently incorporated surrogate or intermediate biomarkers in certain therapeutic development programs. This has accelerated access to innovative treatments but has simultaneously intensified the need for confirmatory clinical evidence and long-term surveillance. The distinction between biological activity, surrogate endpoint improvement, and definitive clinical benefit remains fundamental to scientifically rigorous interpretation of emerging DMD pharmacotherapies.

The contemporary pharmacotherapeutic armamentarium should therefore be understood as an interconnected system rather than a simple list of medications. At one level, pharmacotherapy seeks to preserve muscle function and delay disease progression through corticosteroid and molecularly targeted interventions. At another level, it aims to modify cardiac remodeling, maintain respiratory stability, preserve skeletal integrity, manage endocrine consequences, and alleviate symptoms. These levels of intervention operate simultaneously across different pathophysiological domains and frequently involve medications with markedly different mechanisms and safety profiles. The central pharmacological challenge is consequently to maximize cumulative therapeutic benefit while minimizing cumulative treatment burden.

From an applied perspective, DMD provides an exceptional model for the integration of molecular genetics, RNA biology, medicinal chemistry, epigenetic pharmacology, vector-based therapeutics, immunopharmacology, clinical pharmacology, and pharmacovigilance. The disease demonstrates how understanding a single molecular defect can generate multiple therapeutic strategies operating at different biological levels. It also illustrates the limitations of mechanistic reasoning when molecular correction does not necessarily translate directly into equivalent functional improvement. The distance between molecular target engagement and patient-centered outcomes remains one of the most important scientific considerations in the evaluation of advanced DMD therapeutics.

A comprehensive pharmacological appraisal must consequently evaluate each therapeutic modality through a unified framework encompassing mechanism of action, molecular target, pharmacodynamic effect, pharmacokinetic disposition, genotype dependence, dose–response characteristics and route of administration, therapeutic monitoring, adverse drug reactions, contraindications, drug–drug interactions, organ-specific toxicity, immunological consequences,

and long-term pharmacovigilance. Such an approach is particularly important in DMD because therapeutic exposure may begin during childhood and continue over decades. The consequences of an adverse drug reaction are therefore potentially amplified by cumulative exposure and by the physiological vulnerability associated with progressive neuromuscular deterioration.

The contemporary scientific discourse surrounding DMD pharmacotherapy reflects a transition from conventional symptomatic pharmacology toward molecularly informed, genotype-adapted, multidimensional therapeutic intervention. Corticosteroids remain important because of their established effects on muscle function, whereas exon-skipping agents introduce sequence-specific RNA modulation; givinostat introduces epigenetic pharmacology; and gene replacement introduces vector-mediated transgene expression. These disease-modifying approaches coexist with cardioprotective, respiratory, skeletal, endocrine, analgesic, and other supportive therapies. The resulting therapeutic ecosystem is both scientifically sophisticated and clinically demanding.

The present pharmacological perspective approaches DMD as a complex therapeutic continuum in which molecular correction, pathophysiological modulation, organ protection, symptom control, and safety surveillance must be considered simultaneously. The objective is not merely to catalogue pharmacological agents but to critically contextualize their molecular and cellular mechanisms, pharmacodynamic consequences, pharmacokinetic determinants, therapeutic indications, genotype-specific applicability, adverse-event spectra, toxicological liabilities, and pharmacosafety requirements. Particular emphasis is warranted on the evolving relationship between precision medicine and pharmacotherapy, the interpretation of molecular biomarkers as pharmacodynamic indicators, the management of cumulative treatment toxicity, and the role of pharmacovigilance in defining long-term therapeutic value.

The modernization of DMD pharmacotherapy represents a broader transformation in therapeutic medicine: the movement from empirically selected symptomatic treatment toward mechanistically targeted intervention supported by molecular diagnostics, biomarker-guided assessment, individualized risk stratification, and longitudinal safety surveillance. The future of DMD treatment will likely depend not on a single universally dominant drug but on increasingly rational integration of complementary pharmacological and genetic technologies. Such integration requires rigorous pharmacological characterization, scientifically defensible evidence synthesis, continuous reassessment of benefit–risk relationships, and sustained attention to the practical realities of lifelong treatment. A comprehensive understanding of the contemporary pharmacotherapeutic armamentarium is therefore essential for optimizing therapeutic selection, anticipating toxicity, improving pharmacosafety, and advancing precision-oriented care for individuals affected by Duchenne muscular dystrophy.

BACKGROUND

Duchenne muscular dystrophy (DMD) occupies a distinctive position within contemporary neuromuscular medicine because it represents simultaneously a paradigmatic monogenic disorder, a progressive degenerative myopathy, a multisystem disease, and an increasingly sophisticated target for precision pharmacotherapy. The disorder is inherited in an X-linked pattern and results from pathogenic variation within the DMD gene, which encodes dystrophin, a large cytoskeletal protein indispensable for the structural and functional integrity of striated muscle. The absence or profound deficiency of dystrophin initiates a complex sequence of molecular and cellular abnormalities that progressively compromise skeletal muscle architecture, respiratory function, cardiac performance, and overall physiological resilience. Historically, therapeutic strategies were directed predominantly toward preservation of muscle strength and functional independence through corticosteroid administration and multidisciplinary supportive care. The subsequent emergence of exon-skipping antisense oligonucleotides, epigenetic pharmacotherapy, and adeno-associated-virus-mediated micro-dystrophin gene replacement has

fundamentally altered the therapeutic conceptualization of DMD. Contemporary treatment consequently encompasses interventions acting at different biological levels, from transcriptional and post-transcriptional regulation to intracellular protein restoration, immune modulation, cardiac remodeling, skeletal preservation, and management of treatment-associated toxicity.

Understanding the pharmacotherapeutic evolution of DMD requires recognition that dystrophin deficiency is not an isolated molecular event but the initiating lesion of a progressively amplifying pathophysiological network. Dystrophin is located predominantly beneath the sarcolemma of skeletal and cardiac muscle fibers, where it participates in the dystrophin-associated protein complex and provides a mechanical and signaling interface between intracellular actin and the extracellular matrix. This molecular linkage distributes mechanical forces generated during muscle contraction and protects the sarcolemma from repetitive contraction-induced injury. When dystrophin is absent or markedly reduced, the sarcolemma becomes mechanically unstable. Recurrent membrane disruption permits abnormal intracellular calcium entry and promotes activation of calcium-dependent proteases, mitochondrial injury, reactive oxygen species generation, membrane lipid alteration, and cellular necrosis. Intracellular constituents released from damaged myofibers activate innate immune responses and promote inflammatory signaling. Persistent inflammation subsequently interacts with fibro-adipogenic remodeling, extracellular matrix deposition, and impaired regenerative capacity, progressively converting functional muscle tissue into fibrotic and adipose replacement tissue.

This pathobiological cascade provides a fundamental pharmacological rationale for therapeutic intervention at multiple levels. One strategy is to compensate for the primary absence of dystrophin by restoring production of a structurally or functionally modified dystrophin protein. Another is to modify dystrophin pre-mRNA processing so that the translational reading frame can be restored in genetically appropriate patients. A third approach is to alter downstream pathological pathways, including inflammation, fibrosis, epigenetic dysregulation, and aberrant muscle remodeling. A fourth involves preservation of organs secondarily affected by dystrophin deficiency, particularly the myocardium and respiratory musculature. The contemporary pharmacotherapeutic framework is therefore inherently multidimensional and cannot be adequately represented by a single pharmacological class.

The genetic architecture of DMD has played a decisive role in the transition toward molecularly targeted pharmacotherapy. The considerable size and structural complexity of the *DMD* gene permit numerous pathogenic variant types, including large deletions, duplications, and smaller sequence-level abnormalities. Some variants disrupt the open reading frame and produce premature termination or markedly impaired dystrophin synthesis, whereas others may permit production of truncated dystrophin with variable residual function. The resulting genetic heterogeneity is directly relevant to treatment because some molecular therapies depend upon the precise location and structural consequence of the pathogenic variant. Molecular diagnosis therefore serves not merely as a means of confirming disease but as a therapeutic stratification mechanism.

The concept of genotype-dependent treatment represents a major transition in pharmacological thinking. Traditional pharmacotherapy generally begins with a clinical diagnosis and selects treatment according to disease phenotype. In DMD, by contrast, molecular diagnosis can determine whether a patient is eligible for a particular mechanism-based therapy. Exon-skipping agents provide the clearest example. Their activity depends on whether omission of a specific exon can restore the reading frame of the patient's dystrophin transcript. Thus, pharmacological eligibility is determined by the intersection of drug chemistry and genomic architecture. This model anticipates a broader transformation in medicine in which the molecular substrate of disease increasingly determines therapeutic selection.

Prior to the development of molecularly targeted treatments, glucocorticoids became the principal pharmacological disease-modifying therapy for DMD. Prednisone or prednisolone demonstrated clinically meaningful effects on muscle strength and motor function and became integrated into long-term standards of care. The pharmacological basis of glucocorticoid benefit is complex and incompletely explained by classical anti-inflammatory activity. Glucocorticoids diffuse across cellular membranes and bind intracellular glucocorticoid receptors, generating receptor–ligand complexes that translocate to the nucleus and regulate transcription through glucocorticoid response elements and interactions with other transcriptional regulators. These processes influence inflammatory cytokine production, immune-cell activity, metabolic signaling, muscle protein turnover, and tissue remodeling. In DMD, the resulting systemic effects appear capable of modifying several secondary pathological processes associated with dystrophin deficiency.

The clinical importance of glucocorticoid therapy derives not only from preservation of skeletal muscle performance but also from potential effects on respiratory function and prolongation of ambulation. However, chronic exposure creates a pharmacological paradox: the same systemic biological activity responsible for therapeutic benefit can generate substantial cumulative toxicity. Growth suppression, excessive weight gain, cushingoid features, behavioral disturbances, hypertension, insulin resistance, osteoporosis, vertebral compression fractures, cataracts, infection susceptibility, and hypothalamic–pituitary–adrenal axis suppression may emerge during prolonged treatment. Consequently, glucocorticoid pharmacotherapy illustrates the necessity of balancing pharmacodynamic efficacy against cumulative toxicological burden.

The development of deflazacort expanded the corticosteroid pharmacotherapeutic repertoire. Deflazacort is an oxazoline derivative related structurally to prednisone and undergoes conversion to an active metabolite. Its clinical use in DMD reflects the possibility of modifying the adverse-effect profile of a therapeutic class through structural alteration while maintaining clinically useful glucocorticoid activity. Comparative clinical experience has suggested differences in weight gain and selected metabolic outcomes relative to prednisone, although deflazacort remains associated with important corticosteroid toxicities, including adrenal suppression, skeletal complications, growth effects, and cataract formation. The selection between corticosteroid regimens must therefore be individualized rather than based exclusively on nominal pharmacological potency.

The pharmacology of chronic corticosteroid therapy is further complicated by the developmental context in which treatment is administered. DMD therapy commonly begins during childhood, a period characterized by changing body composition, organ maturation, endocrine development, and evolving pharmacokinetic parameters. Weight-based dosing may therefore require periodic reassessment. Moreover, the adverse consequences of prolonged glucocorticoid exposure can interact with the natural history of DMD, particularly in relation to mobility, bone health, obesity, and cardiometabolic risk. The pharmacological management of steroid toxicity is consequently inseparable from the overall disease-management strategy.

Vamorolone emerged from an alternative approach to corticosteroid medicinal chemistry. Rather than accepting the adverse-effect profile of conventional glucocorticoids as an unavoidable consequence of receptor activation, its development sought to modify the pharmacological characteristics of glucocorticoid-receptor signaling. Vamorolone is structurally distinct from traditional corticosteroids and demonstrates a differentiated receptor pharmacology, with the therapeutic objective of maintaining clinically useful glucocorticoid effects while reducing selected toxicities associated with conventional corticosteroids. Its development exemplifies the increasing importance of structure–activity relationships, receptor pharmacology, and selective modulation of downstream transcriptional signaling in the design of therapies for chronic genetic disorders.

A major conceptual breakthrough occurred with the development of antisense oligonucleotide-mediated exon skipping. The molecular basis of this approach derives from the observation that

selected dystrophin gene deletions can potentially be rendered more compatible with translation by deliberately removing an adjacent exon during pre-mRNA processing. Antisense oligonucleotides are designed to hybridize selectively with complementary sequences within dystrophin pre-mRNA, thereby interfering with recognition of a targeted exon by the splicing machinery. The resulting mature transcript may restore the reading frame and enable synthesis of a shorter dystrophin protein. This strategy does not correct the underlying DNA mutation but modifies RNA processing to produce a more favorable translational outcome.

Eteplirsen, golodirsen, viltolarsen, and casimersen represent major examples of this molecular pharmacological strategy. Their development demonstrates how knowledge of RNA biology can be translated into genotype-specific therapeutics. Eteplirsen targets exon 51, while golodirsen and viltolarsen target exon 53, and casimersen targets exon 45. The precise genetic architecture of the individual patient therefore determines potential eligibility. This therapeutic model differs fundamentally from corticosteroid treatment because the drug does not act primarily through a generalized physiological pathway; rather, its pharmacological activity depends on sequence-specific molecular recognition.

The pharmacodynamics of exon-skipping therapy also requires a layered interpretation. Drug administration produces systemic exposure, followed by tissue distribution and cellular uptake. The pharmacologically relevant event is intracellular interaction with dystrophin pre-mRNA, followed by altered splicing, increased production of the desired transcript, and potentially increased dystrophin protein expression. The ultimate clinical objective is preservation or improvement of muscle function. Thus, a pharmacodynamic continuum extends from molecular target engagement through RNA modification and protein expression to functional phenotype. Each level represents a different biological endpoint, and discordance between these endpoints is possible. A measurable molecular effect does not necessarily predict a proportionate clinical improvement, making endpoint interpretation a central issue in therapeutic evaluation.

Antisense oligonucleotides possess pharmacokinetic properties that distinguish them from traditional small-molecule medicines. Their large molecular size, charge characteristics, protein interactions, tissue distribution, intracellular trafficking, and renal elimination influence systemic disposition. Metabolism may occur through nuclease-mediated pathways rather than classical hepatic cytochrome P450 biotransformation. Repeated intravenous administration is generally required because therapeutic molecular interactions are not necessarily sustained indefinitely. Renal handling is particularly relevant to the safety profile of certain exon-skipping agents, emphasizing the importance of renal monitoring throughout treatment.

The emergence of gene therapy introduced another dimension of molecular pharmacology. Delandistrogene moxeparvovec-rokl employs a recombinant adeno-associated viral vector to deliver a micro-dystrophin transgene. Because the full dystrophin coding sequence is too large for conventional AAV packaging constraints, the therapeutic construct encodes a truncated micro-dystrophin designed to retain critical functional domains. The therapeutic premise is to achieve sustained expression of a dystrophin-derived protein in muscle cells through vector-mediated gene delivery.

The pharmacology of AAV-mediated gene replacement cannot be interpreted using conventional small-molecule principles alone. Following administration, vector particles undergo systemic distribution and tissue tropism, interact with cellular receptors, enter target cells, undergo intracellular processing, and ultimately facilitate delivery of the transgene to the nucleus. Expression of the micro-dystrophin transgene is then influenced by vector biology, cellular environment, immune responses, and tissue-specific factors. The therapeutic effect therefore depends on a chain of biological events rather than a single receptor–drug interaction.

Host immunity constitutes a particularly important determinant of gene-therapy pharmacology. Pre-existing antibodies against the vector may affect eligibility or influence systemic exposure.

Following administration, innate and adaptive immune responses may contribute to hepatic injury, inflammatory reactions, complement activation, or other adverse events. Corticosteroid immunomodulation may therefore be incorporated into the therapeutic protocol. This introduces a complex pharmacological interaction between the gene-therapy platform and immunosuppressive therapy, requiring careful coordination of dosing, monitoring, and risk management.

The liver occupies a central position in gene-therapy safety because vector biodistribution may result in substantial hepatic exposure. Transaminase elevations and other manifestations of hepatocellular injury require structured monitoring, and corticosteroid treatment may be used to mitigate immune-mediated hepatic toxicity. Hematological abnormalities, including thrombocytopenia, and other systemic effects may also occur. The long biological persistence of vector-mediated therapy creates an additional pharmacosafety challenge: unlike conventional drugs that can be rapidly eliminated after discontinuation, the therapeutic transgene may remain biologically active after a single administration. Long-term surveillance is therefore integral to the pharmacological evaluation of gene therapy.

Givinostat represents an additional mechanistic category within contemporary DMD treatment. Its pharmacological target is histone deacetylase activity, placing it within the field of epigenetic pharmacology. Histone acetylation is an important determinant of chromatin organization and transcriptional accessibility, and histone deacetylases participate in the regulation of gene-expression programs involved in muscle differentiation, inflammation, fibrosis, and tissue remodeling. In dystrophic muscle, pathological changes in these pathways may contribute to progressive replacement of functional tissue by fibrotic and adipose components. Pharmacological inhibition of histone deacetylases therefore provides a means of modifying downstream molecular processes without directly correcting the primary *DMD* mutation.

The pharmacodynamic profile of givinostat is consequently distinct from both exon-skipping therapy and gene replacement. Rather than producing a sequence-specific alteration in dystrophin pre-mRNA or introducing a therapeutic transgene, it modifies enzymatic activity within epigenetic regulatory pathways. This illustrates the increasingly heterogeneous nature of DMD pharmacology, in which therapeutic agents can intervene at genomic, transcriptomic, epigenetic, cellular, and systemic physiological levels.

The development of these pharmacological classes has occurred against the background of continuing recognition that DMD is a multisystem disorder. Cardiac involvement is particularly important because dystrophin deficiency affects cardiomyocytes as well as skeletal muscle. Progressive myocardial fibrosis, ventricular dysfunction, conduction abnormalities, and dilated cardiomyopathic remodeling can develop during the disease course. Cardiac pharmacotherapy consequently forms an indispensable component of comprehensive DMD management, even though most cardiovascular agents do not directly modify the primary genetic defect.

Renin–angiotensin-system inhibition has acquired an important role in the pharmacological management of DMD-associated cardiomyopathy. ACE inhibitors and angiotensin-receptor blockers reduce pathological neurohormonal activation, afterload, and myocardial remodeling. Their pharmacodynamic effects include reduction of angiotensin II-mediated vasoconstriction and profibrotic signaling, depending on the specific therapeutic class. Mineralocorticoid-receptor antagonists such as eplerenone can additionally modify aldosterone-mediated myocardial fibrosis and remodeling. B-adrenergic blockers may provide complementary effects by reducing sympathetic drive, heart rate, myocardial oxygen demand, and maladaptive β-adrenergic signaling. These treatments illustrate the importance of early pharmacological cardioprotection in a disorder in which myocardial damage may progress before severe clinical symptoms become apparent.

The safety management of cardiac pharmacotherapy requires attention to renal function, serum electrolytes, systemic blood pressure, cardiac conduction, and treatment combinations. ACE inhibitors, angiotensin-receptor blockers, and mineralocorticoid-receptor antagonists may increase the risk of hyperkalemia and renal dysfunction, particularly when used together or in patients with compromised renal physiology. B-blockers may cause bradycardia, hypotension, fatigue, or conduction abnormalities. Diuretics may become necessary in patients with clinically significant congestion but may produce electrolyte disturbances and intravascular volume depletion. The complexity of this pharmacological environment reinforces the importance of individualized treatment and longitudinal laboratory surveillance.

Respiratory disease constitutes another major determinant of DMD morbidity. Progressive weakness of the diaphragm and accessory respiratory muscles reduces ventilatory reserve and cough effectiveness. Nocturnal hypoventilation may precede overt daytime respiratory failure. Pharmacological treatment has a more limited role than ventilatory support, airway-clearance strategies, and mechanical assistance, but medications may be required to treat superimposed respiratory conditions. Antimicrobial agents, bronchodilators in appropriately selected patients, and mucolytic therapies may be administered according to clinical indication. Their use should nevertheless be differentiated from genuine disease-modifying pharmacotherapy.

Skeletal disease is similarly influenced by the interaction between primary neuromuscular pathology, reduced mechanical loading, immobility, nutritional factors, and chronic glucocorticoid exposure. Low bone mineral density and vertebral compression fractures represent important complications. The pharmacological foundation of skeletal management includes adequate vitamin D and calcium status, nutritional optimization, physical management, and appropriately selected antiresorptive therapy when clinically indicated. Intravenous bisphosphonates have particular relevance for clinically significant vertebral fractures and osteoporosis in DMD. Pharmacological decisions must incorporate skeletal maturity, renal function, biochemical status, fracture history, and cumulative corticosteroid exposure.

The endocrine consequences of long-term corticosteroid treatment introduce another layer of pharmacological complexity. Suppression of the hypothalamic–pituitary–adrenal axis can persist after glucocorticoid discontinuation, creating vulnerability during acute illness, trauma, or surgery. Patients receiving chronic corticosteroids therefore require appropriate recognition of adrenal suppression and implementation of stress-dose strategies when clinically indicated. This is a particularly important example of pharmacological toxicity that may remain clinically silent until a physiological stressor exposes the underlying endocrine deficit.

The long-term management of DMD is therefore characterized by extensive therapeutic exposure across multiple pharmacological classes. Such polypharmacy increases the probability of pharmacodynamic and pharmacokinetic interactions. Agents may share effects on blood pressure, renal function, electrolytes, cardiac conduction, hepatic metabolism, immune function, or central nervous system activity. The potential for cumulative toxicity becomes especially important when disease-specific therapies are combined with corticosteroids and multiple organ-protective medications. Comprehensive medication review is consequently a fundamental component of clinical pharmacology in DMD.

The concept of therapeutic monitoring must also be expanded beyond conventional clinical observation. Molecular therapies require assessment of treatment-specific biomarkers and organ-system safety parameters. Corticosteroid treatment requires surveillance of growth, body mass, blood pressure, glucose metabolism, bone health, ophthalmological status, behavioral effects, and adrenal function. Exon-skipping therapies may require monitoring of renal and infusion-related safety parameters. Gene therapy requires intensive assessment of hepatic, hematological, and immune-mediated toxicity. Givinostat requires laboratory surveillance relevant to its pharmacological effects and potential toxicities. Cardiovascular pharmacotherapy requires serial

evaluation of cardiac structure and function, blood pressure, renal function, and electrolytes. Thus, pharmacosafety is fundamentally a longitudinal process.

The pharmacovigilance implications of DMD therapeutics are particularly significant because many therapeutic agents are administered for prolonged periods, while others represent novel modalities with limited long-term exposure data at the time of initial clinical adoption. Rare-disease populations pose additional challenges for post-marketing surveillance because the number of treated individuals is inherently limited. Adverse drug reactions may therefore be difficult to identify rapidly through conventional statistical signals. Disease registries, longitudinal observational studies, spontaneous reporting systems, electronic health records, and specialized neuromuscular networks provide complementary mechanisms for detecting safety signals.

Another critical background consideration concerns the relationship between pharmacological biomarkers and clinical benefit. DMD therapeutic development increasingly incorporates molecular or biochemical measures such as dystrophin expression, exon-skipping efficiency, or micro-dystrophin production. These biomarkers provide valuable evidence of biological activity, but their interpretation must remain distinct from direct evidence of functional improvement or prolonged survival. The pharmacological chain from molecular target engagement to patient-centered clinical outcome can contain multiple sources of biological variability. Consequently, rigorous evaluation requires integration of molecular, functional, and long-term clinical evidence.

The emergence of precision therapies has also changed the economic and ethical architecture of DMD pharmacotherapy. Genotype-specific therapies may benefit relatively small patient populations and can require sophisticated molecular diagnostics. Gene therapies involve highly complex manufacturing processes and specialized treatment infrastructure. The high technological intensity of these interventions raises questions regarding accessibility, health-system preparedness, treatment prioritization, long-term monitoring, and equitable distribution of advanced therapeutics. These considerations do not diminish the pharmacological value of innovation but demonstrate that therapeutic efficacy exists within a broader clinical and health-system context.

The temporal dimension of treatment is equally important. DMD is progressive, and muscle tissue lost through irreversible fibrosis and adipose replacement cannot necessarily be restored even when molecular dystrophin expression is subsequently achieved. This creates a strong rationale for early therapeutic intervention, particularly before advanced structural damage develops. The pharmacodynamic concept of “earlier is better” may therefore be especially relevant to disease-modifying strategies. Nevertheless, early treatment also implies prolonged exposure and greater cumulative opportunity for adverse events, making pharmacosafety surveillance essential from the beginning of therapy.

The pharmacological development of DMD consequently illustrates a broader transition from conventional disease management to precision therapeutics. The therapeutic target is no longer restricted to symptoms or downstream inflammation. Instead, modern approaches can potentially intervene at the level of RNA processing, epigenetic regulation, transgene expression, or systemic pathophysiological remodeling. This diversification creates opportunities for therapeutic synergy but also demands more sophisticated pharmacological classification and evidence interpretation.

From the perspective of molecular pharmacology, DMD can now be conceptualized as a disease in which multiple therapeutic axes coexist. The first axis involves corticosteroid receptor signaling and modulation of downstream inflammatory and metabolic pathways. The second involves sequence-specific RNA manipulation through antisense oligonucleotides. The third involves epigenetic regulation through histone deacetylase inhibition. The fourth involves vector-mediated genetic delivery and micro-dystrophin expression. The fifth involves organ-protective pharmacotherapy directed toward cardiac and skeletal complications. These axes are not mutually exclusive and may operate concurrently throughout the disease course.

This multidimensional pharmacology also necessitates a distinction between disease-modifying therapy and supportive pharmacotherapy. Disease-modifying interventions seek to alter the trajectory of DMD itself, whereas supportive therapies primarily prevent or treat complications arising from disease progression or its treatment. Cardiovascular drugs, bone-directed therapies, respiratory medications, analgesics, and other symptomatic agents may be essential to survival and quality of life despite not correcting dystrophin deficiency. An academically rigorous pharmacological classification must therefore avoid conflating direct molecular disease modification with management of secondary pathology.

The background to contemporary DMD pharmacotherapy is ultimately characterized by the convergence of genetics, molecular biology, medicinal chemistry, RNA science, epigenetics, vectorology, immunology, and clinical pharmacology. This convergence has transformed DMD from a disorder for which therapeutic options were comparatively limited into a field containing multiple mechanistically distinct disease-modifying strategies. Nevertheless, the existence of several pharmacological modalities does not eliminate uncertainty. Questions remain regarding comparative effectiveness, optimal sequencing, long-term combination strategies, durability of molecular responses, clinically meaningful biomarker thresholds, treatment effects across different disease stages, and the long-term consequences of novel therapeutic platforms.

The concept of pharmacological benefit–risk optimization therefore becomes central. Every therapeutic intervention must be evaluated not only according to whether it produces a measurable biological effect but also according to the magnitude and durability of clinical benefit, treatment burden, route of administration, cumulative toxicity, monitoring requirements, patient-specific risk factors, and alternatives available to the individual patient. In DMD, this assessment is particularly complex because the disease itself is progressive and irreversible, while several treatments are intended for prolonged or lifelong use.

The contemporary pharmacotherapeutic armamentarium thus represents an evolving continuum rather than a static collection of medicines. Conventional corticosteroids continue to provide clinically important benefits; exon-skipping agents provide genotype-specific RNA-directed intervention; givinostat introduces epigenetic modulation; and gene replacement seeks to restore expression of a therapeutically engineered dystrophin derivative. These interventions are embedded within an extensive framework of cardiac, respiratory, skeletal, endocrine, and symptomatic pharmacotherapy. Each therapeutic domain possesses distinct pharmacodynamic objectives, pharmacokinetic determinants, toxicity mechanisms, and pharmacovigilance requirements.

The scientific background of DMD pharmacotherapy supports an integrated approach in which molecular diagnosis, pharmacological classification, mechanism-based therapeutic selection, longitudinal clinical monitoring, and systematic safety assessment are considered components of a single therapeutic continuum. Such an approach is necessary because the pharmacological consequences of DMD extend from the molecular scale of dystrophin gene expression to the organ-system scale of cardiopulmonary failure. The therapeutic challenge is consequently not simply to identify a drug capable of modifying one pathological mechanism but to construct a durable, individualized pharmacological strategy capable of preserving function while minimizing cumulative harm.

The progressive maturation of DMD therapeutics demonstrates the broader potential of precision medicine in inherited disorders. Molecular characterization can determine therapeutic eligibility; molecular pharmacology can provide mechanistically rational interventions; pharmacodynamic biomarkers can assess biological activity; pharmacokinetic science can optimize exposure and safety; and pharmacovigilance can identify adverse events that emerge only during prolonged real-world use. The integration of these disciplines provides the intellectual foundation for contemporary DMD pharmacotherapy.

The comprehensive evaluation of DMD pharmacotherapy must encompass the entire therapeutic continuum, beginning with the molecular consequences of dystrophin deficiency and extending through drug-target interaction, intracellular signaling, tissue distribution, systemic disposition, clinical efficacy, toxicity, and long-term safety. Such an integrated perspective is particularly necessary in an era in which conventional pharmacological agents coexist with RNA therapeutics, epigenetic modulators, and gene-replacement technologies. The expanding therapeutic repertoire offers unprecedented opportunities for altering the clinical trajectory of DMD, while simultaneously imposing increasingly sophisticated requirements for pharmacological characterization and pharmacosafety surveillance.

The contemporary background of DMD pharmacotherapy can therefore be understood as the progressive convergence of disease biology and therapeutic innovation. The historical reliance on corticosteroids established the principle that disease progression could be pharmacologically modified. The subsequent development of exon-skipping agents demonstrated that genetic information could directly guide drug selection. The emergence of histone deacetylase inhibition established downstream epigenetic modulation as another therapeutic strategy, while gene replacement introduced the possibility of durable molecular intervention through vector-mediated transgene delivery. Parallel advances in cardiac, respiratory, skeletal, and endocrine management have transformed supportive pharmacotherapy into an integral determinant of long-term outcomes.

This evolution establishes the scientific rationale for critically examining the contemporary pharmacotherapeutic armamentarium of DMD through an integrated pharmacological framework. Such examination must recognize the heterogeneity of mechanisms, the genotype-specific nature of certain therapies, the complexity of pharmacokinetic and pharmacodynamic relationships, the cumulative burden of chronic treatment, and the distinctive safety challenges associated with advanced therapeutic platforms. Ultimately, the modern pharmacological approach to DMD is defined by the pursuit of precision, mechanistic specificity, therapeutic durability, and individualized benefit–risk optimization. The future trajectory of the field will depend on the continued integration of molecular diagnostics, pharmacology, translational science, clinical evidence, and pharmacovigilance into a coherent therapeutic strategy capable of addressing both the primary genetic pathology and the extensive systemic consequences of dystrophin deficiency.

GOAL

The primary goal of this scientific review is to provide a comprehensive, pharmacologically rigorous evaluation of the contemporary therapeutic armamentarium for Duchenne muscular dystrophy (DMD), integrating drug classification, molecular mechanisms, pharmacodynamics, pharmacokinetics, therapeutic efficacy, adverse drug reactions, toxicological characteristics, pharmacovigilance, and pharmacosafety. Particular emphasis is placed on precision-oriented disease-modifying interventions, genotype-dependent therapeutic selection, long-term benefit–risk optimization, and the rational integration of molecular, cardiopulmonary, skeletal, and symptomatic pharmacotherapy within individualized DMD management.

METHODOLOGY

This review employed a rigorously structured, integrative, and methodologically transparent scientific framework to critically synthesize and interpret contemporary evidence pertaining to the pharmacotherapeutic management of Duchenne muscular dystrophy (DMD). The methodological architecture incorporated systematic evidence identification, critical qualitative appraisal, comparative therapeutic evaluation, and multidimensional pharmacological integration of data concerning both disease-modifying interventions and supportive pharmacotherapeutic strategies.

Particular emphasis was placed on the elucidation of molecular mechanisms of action, pharmacodynamic properties, pharmacokinetic determinants, genotype-dependent therapeutic eligibility, clinical effectiveness, adverse drug reactions, toxicological profiles, clinically significant drug–drug interactions, and comprehensive pharmacosafety considerations. Systematic review and meta-analysis methods synthesized comparative efficacy, safety, pharmacological, and clinically relevant outcome data.

The evidentiary foundation was derived from a conceptually integrated assessment of peer-reviewed biomedical and pharmaceutical literature, contemporary clinical practice guidelines, regulatory and pharmacotherapeutic assessments, established pharmacological and biomedical databases, and publicly accessible prescribing information. Retrieved evidence was critically contextualized according to methodological rigor, clinical relevance, therapeutic endpoint characteristics and duration of follow-up, patient population, genetic eligibility, and level of evidentiary certainty. Pharmacotherapeutic interventions were systematically stratified according to pharmacological class, principal molecular target, mechanism of therapeutic action, clinical indication, therapeutic positioning, translational maturity, and regulatory status. Particular analytical consideration was accorded to corticosteroid-based therapies, exon-skipping antisense oligonucleotides, histone deacetylase inhibition, micro-dystrophin-mediated gene-replacement therapy, and pharmacological approaches directed toward the prevention or management of cardiac, respiratory, musculoskeletal, skeletal, endocrine, and other clinically consequential manifestations of DMD.

The analytical framework was further informed by a multidisciplinary precision-medicine paradigm integrating molecular genetics, pharmacology, clinical pharmacology, pharmacokinetics, pharmacodynamics, toxicology, pharmacovigilance, therapeutics, and evidence-based medicine. Therapeutic benefit was interpreted within a longitudinal benefit–risk framework that considered not only efficacy-related outcomes but also treatment burden, tolerability, cumulative toxicity, clinically meaningful safety signals, pharmacological limitations, genotype-specific applicability, and uncertainties surrounding long-term therapeutic effectiveness. Where appropriate, quantitative findings from eligible comparative studies and meta-analytic evidence were interpreted in relation to the heterogeneity of study populations, intervention characteristics, comparator regimens, outcome measures, follow-up periods, and methodological quality.

Throughout the analytical process, principles of scientific integrity, methodological transparency and reproducibility, balanced interpretation of therapeutic benefits and risks, and avoidance of unsupported clinical or causal inferences were rigorously maintained. Particular methodological distinction was made between established standard-of-care interventions, conditionally authorized or accelerated-approval therapies, emerging therapeutic modalities, and investigational pharmacological strategies. This distinction was considered essential for accurately characterizing the contemporary DMD pharmacotherapeutic evidence base while minimizing interpretative bias and preserving an appropriate separation between demonstrated clinical efficacy, surrogate or molecular endpoints, regulatory authorization, and preliminary translational evidence.

RESULTS AND DISCUSSION

Overview of Therapeutic Categories

The contemporary DMD pharmacotherapeutic armamentarium comprises five principal mechanistic categories: (1) glucocorticoid receptor modulators (corticosteroids); (2) RNA-directed molecular therapeutics (exon-skipping antisense oligonucleotides); (3) epigenetic modulators (histone deacetylase inhibitors); (4) gene-replacement therapy (AAV-mediated micro-dystrophin delivery); and (5) supportive pharmacotherapy (cardiac, respiratory, skeletal, endocrine management). (See **Table 1).** Summarizes the key characteristics of major therapeutic classes.

Table 1. Overview of Contemporary DMD Pharmacotherapeutic Classes

Therapeutic Class	Mechanism of Action	Route	FDA Approval	Key Agents	Primary Indication
Glucocorticoids	Glucocorticoid receptor modulation	Oral	1960s-2023	Prednisone, Deflazacort, Vamorolone	Disease modification
Exon-skipping ASOs	Pre-mRNA splicing modulation	IV	2016-2021	Eteplirsen, Golodirsen, Viltolarsen, Casimersen	Genotype-specific restoration
HDAC Inhibitors	Epigenetic regulation	Oral	2024	Givinostat	Disease modification
Gene Therapy	AAV-mediated micro-dystrophin delivery	IV	2023	Delandistrogene moxeparvovec	Protein replacement
Supportive	Various	Various	Variable	ACEi, ARB, β-blockers, bisphosphonates	Complication management

Integrated Interpretation of the Contemporary Pharmacotherapeutic Evidence

The synthesis of contemporary pharmacological evidence demonstrates that the therapeutic paradigm of Duchenne muscular dystrophy (DMD) has undergone a fundamental transformation from predominantly supportive management toward a multidimensional pharmacotherapeutic model incorporating disease modification, molecular intervention, genetic correction, epigenetic regulation, organ protection, and individualized long-term risk management. The pharmacotherapeutic armamentarium can no longer be conceptualized as a homogeneous collection of drugs directed toward muscle weakness. Rather, it represents a hierarchical therapeutic system operating at multiple biological levels, ranging from modulation of glucocorticoid receptor signaling and transcriptional activity to sequence-specific modification of dystrophin pre-mRNA, delivery of micro-dystrophin genetic constructs, inhibition of histone deacetylases, and pharmacological preservation of cardiac, respiratory, skeletal, and endocrine function.

The principal finding emerging from the integrated analysis is that DMD pharmacotherapy is intrinsically genotype-, phenotype-, age-, ambulatory-status-, organ-function-, and treatment-history dependent. This heterogeneity is particularly important because different therapeutic modalities address different pathobiological components of the disease. Corticosteroids remain foundational systemic therapies because of their broad effects on muscle inflammation, strength,

motor function, and disease progression. However, they do not correct the underlying *DMD* genetic defect. Exon-skipping antisense oligonucleotides (ASOs), in contrast, are mechanistically dependent upon specific mutation classes and aim to restore the translational reading frame by modifying pre-mRNA splicing. Givinostat operates through epigenetic modulation and inflammatory/fibrotic pathways and therefore has a broader genotype-independent conceptual applicability. Gene replacement therapy represents a fundamentally different pharmacological paradigm, introducing a micro-dystrophin transgene through an adeno-associated viral vector. Consequently, contemporary DMD treatment reflects the convergence of conventional pharmacology, molecular therapeutics, RNA pharmacology, gene therapy, immunopharmacology, and precision medicine.

This multidimensional therapeutic structure is consistent with modern DMD care recommendations emphasizing coordinated management of neuromuscular, cardiac, respiratory, skeletal, endocrine, orthopaedic, nutritional, and psychosocial manifestations. Contemporary care has also contributed to increased survival and increasingly prolonged transition into adulthood, making long-term pharmacological safety and cumulative treatment burden progressively more important.

Corticosteroids as the Foundational Pharmacotherapeutic Intervention

The evidence synthesis confirms that glucocorticoid therapy remains a central component of DMD pharmacotherapy. Prednisone/prednisolone and deflazacort established the pharmacological foundation upon which subsequent disease-modifying therapies have been developed. Their therapeutic value derives not from restoration of dystrophin but from modulation of inflammatory signaling, transcriptional networks, muscle membrane stability, protein turnover, immune responses, and potentially regenerative processes.

Prednisone exerts its pharmacodynamic effects principally through intracellular glucocorticoid receptor activation. Following cellular entry, the steroid binds the cytoplasmic glucocorticoid receptor, inducing conformational rearrangement, dissociation of chaperone proteins, nuclear translocation, and interaction with glucocorticoid-responsive regulatory elements and transcription factors. This process modifies expression of multiple genes involved in inflammatory signaling, cytokine production, immune-cell activity, cellular metabolism, and tissue remodeling. The pharmacological effect in DMD therefore represents a complex network phenomenon rather than a single molecular target.

Clinical evidence has consistently supported corticosteroids as therapeutically meaningful agents. Historical guideline evidence demonstrated improvements in muscle strength and pulmonary function with prednisone, while deflazacort has demonstrated beneficial effects on strength, timed motor performance, loss of ambulation, and several longer-term disease outcomes. The American Academy of Neurology evidence synthesis identified prednisone as a treatment that should be offered for strength and pulmonary function, while deflazacort may be offered for strength, timed motor function, and delaying loss of ambulation.

The pharmacological advantage of corticosteroids is therefore broad biological activity. Their principal limitation is equally broad systemic toxicity. Chronic exposure may result in excessive weight gain, Cushingoid features, growth suppression, behavioral and psychiatric effects, hypertension, glucose intolerance, osteoporosis, vertebral fractures, cataracts, infection susceptibility, adrenal suppression, and other metabolic complications. Consequently, the clinical value of corticosteroids must be understood through a longitudinal benefit–risk framework rather than through short-term motor outcomes alone.

Prednisone and deflazacort demonstrate partially overlapping therapeutic effects but distinct adverse-effect profiles. Prednisone is commonly associated with greater weight gain, whereas deflazacort has been associated with a comparatively higher risk of cataracts. These differences

are pharmacologically relevant because chronic DMD therapy frequently begins during childhood and may continue for many years. Thus, the optimal corticosteroid cannot be selected solely according to efficacy; body composition, growth trajectory, bone health, ophthalmological status, behavioral vulnerability, metabolic risk, family preferences, adherence, and access must also be considered.

An important pharmacosafety finding is that chronic corticosteroid exposure produces a state of hypothalamic–pituitary–adrenal axis suppression. Abrupt discontinuation is therefore potentially hazardous. Intercurrent infection, trauma, surgery, and other physiological stressors can precipitate adrenal insufficiency if glucocorticoid requirements are not appropriately adjusted. The pharmacological management of DMD consequently requires an emergency steroid strategy and appropriate stress-dose planning. This illustrates an important principle of DMD pharmacotherapy: adverse-effect prevention cannot be separated from therapeutic continuity.

Vamorolone and the Emergence of Dissociated Glucocorticoid Pharmacology

The introduction of vamorolone represents a significant conceptual development in corticosteroid pharmacology. Unlike conventional glucocorticoids, vamorolone was developed with the objective of retaining clinically relevant glucocorticoid-associated activity while modifying receptor interactions associated with selected corticosteroid toxicities. Its pharmacology therefore illustrates the emerging concept of receptor-selective or functionally differentiated steroid signaling.

Vamorolone received FDA approval in 2023 for treatment of DMD in patients two years of age and older. The regulatory evidence included a controlled clinical trial demonstrating improvement in the prespecified motor endpoint of time-to-stand velocity compared with placebo. The recommended U.S. regimen is weight based, administered orally once daily, with dosing individualized according to body weight and tolerability.

The pharmacodynamic importance of vamorolone lies in its altered interaction with the glucocorticoid receptor and related steroid receptor signaling. Its clinical development reflects the proposition that corticosteroid efficacy and corticosteroid toxicity may not be inseparable pharmacological phenomena. Nevertheless, the distinction should not be interpreted as evidence of complete elimination of steroid-related toxicity. Vamorolone remains a corticosteroid and can produce clinically important endocrine, metabolic, immune, and adrenal effects. Therefore, its clinical positioning should be understood as an alternative corticosteroid strategy rather than as a completely toxicity-free replacement for conventional glucocorticoids. From a pharmacotherapeutic perspective, vamorolone expands individualized corticosteroid selection. A patient experiencing clinically consequential weight gain, growth suppression, behavioral effects, or other complications under prednisone or deflazacort may represent a candidate for an alternative steroid strategy, subject to individualized assessment and local regulatory guidance. The significance of vamorolone is therefore both therapeutic and conceptual: it demonstrates that modification of receptor pharmacology can potentially reshape the efficacy–toxicity relationship of long-term steroid treatment.

Corticosteroid Adverse Effects

Chronic corticosteroid exposure produces substantial systemic toxicity, including growth suppression, excessive weight gain, Cushingoid features, behavioral disturbances, hypertension, insulin resistance, osteoporosis, vertebral compression fractures, cataracts, infection susceptibility, and hypothalamic–pituitary–adrenal axis suppression. Prednisone is associated with comparatively greater weight gain, whereas deflazacort has been associated with higher cataract risk. Vamorolone has been developed to modify this efficacy–toxicity relationship and appears to

preserve glucocorticoid-associated motor benefit while potentially altering selected adverse-effect profiles. See Table-2.

Table 2. Comparative Corticosteroid Evidence

Parameter	Prednisone/Prednisolone	Deflazacort	Vamorolone
Evidence Maturity	Very High (decades)	High	Moderate (emerging)
Functional Efficacy	Strong	Strong	Strong
6MWD Preservation	Reference	+28.3 to +39.54 m	Data available
Timed Motor Improvement	Yes	Yes	Yes
Weight Gain	High	Moderate	Reduced
Cataract Risk	Moderate	Higher	Under evaluation
Growth Effects	Significant	Significant	Possibly reduced
Adrenal Suppression	Significant	Significant	Requires monitoring
Route	Oral	Oral	Oral
FDA Approval	Established	Established	2023

Exon-Skipping Antisense Oligonucleotide Therapy

The development of exon-skipping ASOs constitutes one of the clearest examples of precision pharmacology in DMD. Eteplirsen, golodirsen, viltolarsen, and casimersen target specific dystrophin pre-mRNA sequences and are applicable only to genetically defined patient populations. Their pharmacological objective is not conventional receptor modulation but sequence-specific manipulation of RNA processing.

Eteplirsen is designed to induce skipping of exon 51. Golodirsen and viltolarsen target exon 53, whereas casimersen targets exon 45. By binding complementary sequences in dystrophin pre-mRNA, these oligonucleotides interfere with normal splice-site recognition and promote exclusion of the targeted exon. In patients with suitable deletions, this can restore the translational reading frame and permit synthesis of a shorter but potentially functional dystrophin protein. This mechanism represents a highly individualized pharmacological intervention. Therapeutic eligibility is therefore inseparable from molecular genetic characterization. A patient cannot be considered an appropriate candidate merely because DMD has been clinically diagnosed. The specific pathogenic variant must be characterized and demonstrated to be amenable to the relevant exon-skipping strategy.

The pharmacodynamic endpoint of ASO therapy is consequently situated at an intermediate biological level. The drug first produces molecular target engagement and altered RNA splicing;

this should increase dystrophin production; increased dystrophin should improve sarcolemmal stability; and improved membrane integrity should ultimately contribute to preservation of muscle function. Each step introduces potential biological attenuation. Therefore, increased dystrophin expression should not automatically be equated with proportional clinical improvement. This distinction is particularly important for interpretation of regulatory evidence. The FDA continues to identify several exon-skipping therapies within the accelerated-approval framework, with confirmatory evidence requirements intended to verify clinical benefit. The scientific interpretation should consequently distinguish molecular surrogate effects from demonstrated patient-centered outcomes.

The pharmacokinetic properties of ASOs also differ fundamentally from those of conventional small-molecule drugs. These compounds are generally administered parenterally, distribute predominantly through extracellular and intracellular pathways characteristic of oligonucleotide therapeutics, and undergo tissue uptake followed by intracellular processing and degradation rather than conventional hepatic CYP-mediated metabolism. Renal elimination and renal tubular exposure are particularly relevant to safety surveillance. This is clinically important because repeated administration creates cumulative exposure over many years.

Among adverse reactions, infusion-related reactions and renal safety considerations are particularly relevant. Appropriate renal monitoring is therefore an integral component of pharmacosafety. The risk–benefit assessment must also account for repeated intravenous administration, treatment burden, venous access, adherence, and the uncertainty surrounding the magnitude of long-term clinical benefit. The four exon-skipping agents collectively demonstrate an important principle of precision pharmacotherapy: genetic specificity increases mechanistic precision but simultaneously restricts the eligible population. A treatment capable of targeting one exon cannot automatically be generalized to patients carrying other mutation classes. Consequently, genetic diagnosis becomes not merely diagnostic but pharmacologically determinative.

Givinostat and Epigenetic Pharmacology

Givinostat represents a distinct pharmacological category within DMD therapy. FDA approval in March 2024 established it as a histone deacetylase inhibitor for patients six years of age and older. Importantly, it was described by the FDA as the first nonsteroidal treatment approved for DMD across genetic variants.

The mechanistic rationale for HDAC inhibition derives from the recognition that DMD pathophysiology extends beyond dystrophin deficiency itself. Chronic muscle injury activates inflammatory pathways, alters transcriptional regulation, stimulates fibro-adipogenic remodeling, and disrupts regenerative homeostasis. Histone acetylation and deacetylation influence chromatin accessibility and transcriptional activity. By inhibiting HDAC activity, givinostat modifies the expression of multiple genes involved in muscle remodeling and inflammation.

This pharmacology distinguishes givinostat from exon-skipping therapies. Whereas ASOs operate primarily through sequence-specific RNA processing, givinostat acts through epigenetic regulation and downstream pathophysiological pathways. Consequently, its therapeutic applicability is not restricted to a single exon-specific mutation. This provides an important expansion of pharmacological coverage.

The safety profile of givinostat nevertheless requires structured laboratory surveillance. FDA prescribing information specifies baseline platelet counts and triglyceride assessment and identifies dose modification requirements associated with thrombocytopenia, diarrhea, hypertriglyceridemia, and QTc prolongation. These findings illustrate how pharmacosafety must be incorporated into the therapeutic design itself rather than treated as a secondary post-prescription issue. Givinostat is therefore particularly instructive from a clinical-pharmacology

perspective because it requires integration of pharmacodynamics with laboratory pharmacovigilance. Platelet monitoring, lipid assessment, gastrointestinal surveillance, electrocardiographic considerations, and assessment of interacting medications are not optional adjuncts but components of responsible administration.

Gene-Replacement Therapy and the Pharmacology of Delandistrogene Moxeparvovec-Rokl

Gene replacement represents the most technologically complex component of the contemporary DMD pharmacotherapeutic armamentarium. Delandistrogene moxeparvovec-rokl uses an AAVrh74 vector to deliver a micro-dystrophin transgene. Its mechanism therefore differs fundamentally from conventional pharmacotherapy and RNA-based exon skipping.

The therapeutic concept involves vector-mediated delivery of genetic material to muscle cells, followed by intracellular processing and expression of the micro-dystrophin construct. Because the native dystrophin gene is exceptionally large, the therapeutic construct cannot reproduce the entire endogenous protein. Instead, a truncated micro-dystrophin is engineered to retain selected functional domains capable of supporting aspects of sarcolemmal integrity. The theoretical pharmacodynamic advantage is substantial: rather than transiently modifying RNA splicing or suppressing inflammatory pathways, gene therapy seeks to establish persistent intracellular expression of a therapeutic protein. However, this potential durability introduces equally substantial biological complexity. The pharmacology of an AAV-based gene therapy cannot be fully characterized using conventional plasma concentration measurements. The relevant exposure encompasses vector biodistribution, cellular uptake, transduction, intracellular persistence, transgene expression, immune recognition, and tissue-specific effects. Current regulatory information demonstrates why this modality requires particularly rigorous pharmacosafety assessment. As of the current FDA labeling, Elevidys is indicated for ambulatory patients aged four years and older with confirmed DMD mutation, and the labeling includes major limitations relating to hepatic risk.

The evolution of the Elevidys safety profile is one of the most important pharmacovigilance findings in contemporary DMD therapy. In 2025, the FDA reported fatal acute liver failure in non-ambulatory pediatric patients following treatment and subsequently required a boxed warning and revised the indication to restrict use to ambulatory patients aged four years and older. This development illustrates the dynamic nature of pharmacosafety knowledge in advanced therapies. Pre-approval clinical trials may not fully characterize rare, severe, or phenotype-dependent adverse events because exposure populations are comparatively small and follow-up is necessarily limited.

The Elevidys experience therefore demonstrates that regulatory approval does not represent the endpoint of evidence generation. For complex gene therapies, post-marketing surveillance becomes an essential extension of the clinical trial evidence base. Acute liver injury, transaminase elevation, immune-mediated responses, thrombocytopenia, and other systemic effects require active monitoring. Pre-treatment assessment for relevant hepatic abnormalities and infectious conditions is consequently fundamental to patient selection.

The therapeutic risk is particularly difficult to conceptualize because the intervention may be effectively irreversible at the level of vector administration. Conventional drugs can often be discontinued when toxicity occurs. Gene therapy presents a different pharmacological reality: once vector-mediated transduction has occurred, treatment cannot simply be “withdrawn” in the conventional sense. This makes patient selection, informed consent, pre-treatment screening, immunological evaluation, corticosteroid management, and long-term surveillance ethically and clinically indispensable.

Comparative Pharmacological Architecture of Current Disease-Modifying Therapies

The results demonstrate that contemporary DMD therapies can be conceptually organized into four principal mechanistic categories: receptor-mediated anti-inflammatory modulation, RNA-directed molecular intervention, epigenetic pharmacology, and gene replacement. Conventional corticosteroids operate primarily through glucocorticoid receptor signaling and exert broad systemic effects. Their principal advantage is extensive clinical experience, broad applicability, oral administration, and relatively low acquisition complexity compared with advanced molecular therapies. Their principal limitation is cumulative systemic toxicity.

Vamorolone extends this category by attempting to modify the pharmacological relationship between glucocorticoid activity and adverse effects. Its role is therefore particularly relevant to individualized steroid selection. Exon-skipping ASOs provide mutation-specific molecular intervention. Their principal advantage is mechanistic precision; their principal limitations include restricted genetic eligibility, repeated intravenous administration, uncertain magnitude of long-term clinical benefit for some products, and treatment burden.

Givinostat provides genotype-independent epigenetic modulation for patients meeting its age criteria. Its pharmacological advantages include oral administration and broader genetic applicability, whereas laboratory and electrophysiological monitoring remain important limitations.

Gene replacement offers the possibility of durable micro-dystrophin expression but introduces the greatest complexity regarding immune response, hepatic injury, vector-related toxicity, irreversibility, eligibility, and long-term uncertainty.

The resulting therapeutic architecture indicates that there is no universally superior DMD drug. Instead, treatment selection must be conceptualized as an optimization problem involving genotype, age, ambulatory status, disease stage, prior steroid exposure, cardiac and respiratory phenotype, hepatic and renal function, immunological status, treatment access, adherence, patient/family values, and the evidence strength associated with each therapeutic modality.

Cardiac Pharmacotherapy as a Disease-Modifying Supportive Domain

The analysis further demonstrates that DMD pharmacotherapy cannot be restricted to skeletal muscle. Cardiomyopathy is a major determinant of morbidity and mortality, and pharmacological cardiac protection is therefore an essential component of comprehensive treatment.

Dystrophin deficiency in cardiomyocytes results in progressive myocardial injury, fibrosis, ventricular dysfunction, arrhythmogenic substrate formation, and eventual heart failure. The cardiac pharmacotherapeutic objective is consequently to delay remodeling and preserve ventricular function rather than simply treat symptomatic heart failure after advanced dysfunction has developed.

Renin–angiotensin system inhibition, including ACE inhibitors and angiotensin receptor blockers, has an established role in DMD cardiac management. Beta-adrenergic blockade can be incorporated according to cardiac phenotype and ventricular function, while mineralocorticoid receptor antagonists such as eplerenone or spironolactone provide additional antifibrotic and cardioprotective strategies.

The rationale for early pharmacological intervention is particularly strong because myocardial fibrosis may develop before substantial reduction in conventional measures of ventricular systolic function. Therefore, pharmacological surveillance should not be based exclusively on overt heart failure symptoms. Serial cardiac imaging and appropriately timed pharmacological intervention are integral to contemporary DMD care. Consensus-based cardiac management emphasizes surveillance and pharmacological strategies intended to slow the development and progression of cardiomyopathy.

Cardiac pharmacotherapy also demonstrates the importance of interaction between disease-modifying and supportive therapies. Chronic corticosteroid treatment may influence blood pressure, glucose metabolism, and cardiovascular risk, whereas ACE inhibitors, ARBs, beta-blockers, and mineralocorticoid receptor antagonists can produce hypotension, renal dysfunction, or electrolyte disturbances. Consequently, the pharmacological profile of the entire regimen must be considered rather than evaluating each drug in isolation.

Respiratory Pharmacotherapy and the Primacy of Non-Pharmacological Support

Respiratory decline in DMD is predominantly attributable to progressive respiratory muscle weakness, impaired cough, nocturnal hypoventilation, and eventual ventilatory failure. The evidence synthesis demonstrates that pharmacological therapy has a more limited role in this domain than in skeletal muscle or cardiac disease.

Bronchodilators may be appropriate when concomitant reactive airway disease or bronchospasm exists, but they do not reverse the primary respiratory muscle pathology of DMD. Similarly, mucolytic therapy may occasionally assist secretion management but does not replace mechanical airway-clearance strategies.

The most consequential respiratory interventions are frequently non-pharmacological, including lung-volume recruitment, assisted coughing, and non-invasive ventilatory support. Contemporary respiratory recommendations describe a progressive strategy involving these modalities, with ventilation introduced initially during sleep and expanded as respiratory failure progresses. This finding is important because it prevents pharmacological overextension. A comprehensive pharmacotherapeutic review should explicitly distinguish therapies that modify the disease from interventions that manage complications. Attempting to classify every respiratory intervention as a “DMD drug” would obscure the actual pharmacological hierarchy.

Bone, Endocrine, and Metabolic Pharmacotherapy

Long-term glucocorticoid treatment substantially modifies skeletal and endocrine physiology. DMD patients are vulnerable to reduced bone mineral density, vertebral compression fractures, delayed growth, pubertal abnormalities, and endocrine disturbances. The therapeutic response must therefore include monitoring and targeted pharmacotherapy.

Vitamin D and calcium optimization are foundational but should be guided by nutritional status, laboratory evaluation, and individualized requirements rather than indiscriminate supplementation. In patients with clinically significant osteoporosis or vertebral fractures, intravenous bisphosphonates may be considered according to specialist assessment and pediatric bone-health protocols. The interaction between corticosteroid exposures, reduced mobility, altered mechanical loading, endocrine dysfunction, and nutritional status means that bone pharmacotherapy should not be considered independently of the neuromuscular disease. Reduced ambulation decreases skeletal loading, while glucocorticoids interfere with osteoblast function, calcium homeostasis, and bone remodeling. Thus, pharmacological preservation of bone health represents a secondary but essential component of DMD pharmacosafety.

The risk of vertebral fracture is particularly significant because clinically silent fractures may progress before recognition. Pharmacotherapeutic decisions should therefore be integrated with radiographic and densitometric surveillance, endocrine assessment, vitamin D status, and functional evaluation.

Pharmacokinetic Complexity and the Challenge of Long-Term Pediatric Therapy

One of the most important findings of the integrated pharmacological analysis is that DMD treatment occurs within a uniquely complex pharmacokinetic environment. Many patients begin

therapy during early childhood, undergo substantial changes in body mass and body composition, and may experience altered renal, hepatic, nutritional, and physiological parameters over time.
For conventional small molecules such as corticosteroids, pharmacokinetic parameters include absorption, systemic exposure, protein binding, hepatic metabolism, elimination, and receptor-mediated pharmacodynamics. Weight-based dosing can therefore require regular reassessment. For ASOs, conventional small-molecule pharmacokinetic assumptions are less applicable. Their distribution and intracellular trafficking, tissue uptake, degradation, and renal elimination produce distinct exposure–response relationships. The relationship between plasma exposure and muscle dystrophin production may not be linear.
Gene therapy represents an even greater departure from conventional pharmacokinetic paradigms. Traditional plasma concentration–time curves are insufficient to describe therapeutic exposure. Vector biodistribution, transduction efficiency, tissue tropism, immune neutralization, intracellular persistence, and transgene expression become central pharmacological variables. These differences indicate that DMD pharmacology requires a broadened definition of pharmacokinetics. The field increasingly encompasses not only absorption and elimination but also molecular biodistribution, cellular uptake, intracellular persistence, target engagement, biomarker modulation, and biological durability.

Pharmacodynamic Biomarkers Versus Clinical Outcomes

A major interpretive challenge identified in the evidence base is the distinction between molecular pharmacodynamic effects and clinically meaningful patient outcomes. Dystrophin restoration is biologically compelling, but increased dystrophin expression does not necessarily translate proportionally into preservation of ambulation, respiratory function, cardiac function, quality of life, or survival. The distinction is particularly relevant to molecular therapies developed under accelerated regulatory pathways. A drug may demonstrate target engagement or biomarker improvement while the magnitude of clinical benefit remains incompletely established. FDA accelerated-approval records for several exon-skipping products continue to identify confirmatory studies intended to verify clinical benefit. This creates a hierarchy of evidence. At the molecular level, target engagement and protein restoration are important. At the intermediate physiological level, changes in muscle function, strength, and timed motor performance are more directly patient relevant. At the highest clinical level, preservation of ambulation, respiratory independence, cardiac function, survival, and quality of life provide the most consequential endpoints.
A scientifically rigorous interpretation therefore requires avoiding the assumption that surrogate improvement automatically constitutes equivalent clinical benefit. This is particularly important when comparing mechanistically diverse therapies.

Adverse Drug Reactions and Toxicological Profiles

The integrated results indicate that DMD pharmacotherapy produces a heterogeneous spectrum of adverse effects that correspond closely to therapeutic mechanism.
For corticosteroids, toxicity is predominantly systemic and cumulative, including metabolic, endocrine, skeletal, ophthalmological, psychiatric, immunological, and cardiovascular effects. The major pharmacosafety challenge is therefore long-term exposure management. For exon-skipping ASOs, the principal safety considerations involve repeated intravenous administration, infusion reactions, renal effects, and treatment burden. Their toxicity profile is mechanistically different from glucocorticoid toxicity and requires laboratory monitoring tailored to the individual agent.

For givinostat, thrombocytopenia, gastrointestinal effects, hypertriglyceridemia, and QTc prolongation represent particularly important safety domains, necessitating baseline and longitudinal monitoring.
For gene replacement therapy, toxicity may involve immune activation, hepatic injury, hematological abnormalities, and potentially severe acute liver failure. The post-marketing Elevidys experience demonstrates that rare catastrophic toxicity can emerge only after broader clinical exposure. The FDA's subsequent boxed warning illustrates the importance of continuous pharmacovigilance for advanced therapeutic products.
The comparative toxicological analysis therefore demonstrates that no therapeutic modality is intrinsically “safe” or “unsafe.” Safety is conditional upon mechanism, patient phenotype, dose, exposure duration, concomitant treatment, monitoring intensity, and biological susceptibility.

Pharmacovigilance as an Essential Component of DMD Pharmacotherapy

The findings strongly support the proposition that pharmacovigilance should be considered a structural component of DMD treatment rather than a post hoc regulatory activity. This is particularly relevant because DMD is a rare disease, clinical trials frequently involve relatively small populations, and many therapies are administered over prolonged periods.
The limitations of rare-disease trials create a substantial evidentiary problem. Pre-approval studies may be adequately powered to detect changes in selected endpoints but may be insufficient to identify rare catastrophic adverse events. This limitation is particularly important for gene therapies, where biological exposure may be highly heterogeneous.
The Elevidys experience is an exemplary demonstration of this phenomenon. Initial regulatory authorization was followed by post-marketing reports of fatal acute liver failure in non-ambulatory patients, subsequent FDA investigation, and eventual regulatory revision with a boxed warning and restriction of the indicated population. This sequence illustrates a central pharmacovigilance principle: the risk profile of a medicine is not static. Regulatory knowledge evolves as exposure increases and as previously unrecognized patient-specific vulnerabilities become apparent.
Pharmacovigilance in DMD should therefore incorporate spontaneous adverse-event reporting, longitudinal registries, laboratory surveillance, standardized clinical outcome measurement, cardiac and respiratory monitoring, genetic stratification, and long-term follow-up. Real-world evidence is particularly valuable because it can identify interactions between therapeutic exposure and phenotypic variables that may be underrepresented in controlled trials.

Polypharmacy and Drug–Drug Interaction Considerations

As survival increases and disease management becomes more comprehensive, polypharmacy becomes increasingly relevant. A patient may simultaneously receive corticosteroids, molecular disease-modifying therapy, cardiac medications, bone-directed interventions, analgesics, nutritional supplementation, antibiotics, anesthetic agents, and other supportive treatments.
The interaction burden is therefore multidimensional. Pharmacokinetic interactions may arise through effects on metabolism or elimination, whereas pharmacodynamic interactions may involve overlapping effects on blood pressure, renal function, electrolytes, QT interval, immune function, hepatic integrity, or bone metabolism.
Givinostat provides an example in which interaction assessment is clinically relevant because QTc prolongation and laboratory abnormalities may be amplified by concomitant medications or pre-existing conditions. Corticosteroids introduce another dimension because chronic exposure modifies metabolic and endocrine physiology and may interact indirectly with other treatments.
Gene therapy introduces immunopharmacological interactions of a different nature. Corticosteroids may be used to modulate immune responses surrounding vector administration, meaning that immunosuppression itself becomes part of the therapeutic protocol. Consequently,

corticosteroid exposure in gene-therapy patients cannot always be interpreted simply as conventional DMD maintenance therapy.

Precision Pharmacotherapy and Genotype–Phenotype Integration

The results demonstrate that DMD represents an exemplary model of precision medicine. The molecular diagnosis increasingly determines therapeutic eligibility, particularly for exon-skipping strategies. The genotype must therefore be interpreted not merely as an etiological descriptor but as a pharmacological biomarker. However, genotype alone is insufficient. Two individuals with the same mutation may demonstrate different rates of disease progression because of modifier genes, environmental factors, treatment history and age at treatment initiation, corticosteroid exposure, adherence, comorbidities, and other biological variables.

The ideal precision-pharmacotherapy framework therefore requires integration of genotype, phenotype, molecular biomarkers, clinical trajectory, organ-system status, and treatment tolerability. This multidimensional model is superior to a mutation-only approach because the therapeutic objective is not simply molecular correction but preservation of meaningful function and quality of life. The concept of precision pharmacotherapy consequently extends beyond selecting an exon-skipping drug. It encompasses corticosteroid selection, treatment sequencing, cardiac prophylaxis, bone protection, respiratory surveillance, nutritional management, and individualized pharmacosafety.

Therapeutic Sequencing and Combination Strategies

The contemporary evidence base increasingly supports consideration of therapeutic sequencing rather than isolated drug selection. Corticosteroids may remain foundational while molecular or epigenetic therapies are introduced. Cardiac and respiratory interventions may proceed in parallel rather than sequentially. Bone protection may become necessary as cumulative steroid exposure increases.

However, combination therapy introduces unresolved questions. The pharmacodynamic interaction between chronic glucocorticoids and newer disease-modifying therapies may be beneficial, neutral, or potentially problematic depending upon the mechanism. For example, corticosteroids may remain clinically necessary in patients receiving molecular therapies, while the cumulative adverse-effect burden requires ongoing reassessment.

Combination strategies therefore require evidence that extends beyond the efficacy of individual agents. Important research questions include whether early molecular intervention modifies the required duration or intensity of corticosteroid treatment, whether different disease-modifying therapies can be safely combined, whether combination treatment produces additive or redundant biological effects, and whether biomarker responses can guide treatment de-escalation. At present, these issues require cautious interpretation. Absence of evidence should not be interpreted as evidence of therapeutic equivalence or superiority.

Ethical Dimensions of Advanced DMD Pharmacotherapy

The ethical analysis of DMD pharmacotherapy is inseparable from the clinical pharmacology of rare disease. Patients and families must make decisions involving therapies that may differ substantially in evidence maturity, cost and route of administration, eligibility, uncertainty, reversibility, and toxicity.

The principle of informed consent is particularly important for advanced therapies. Patients and caregivers should understand not only the potential therapeutic benefits but also the degree of evidentiary uncertainty, the distinction between biomarker effects and established clinical outcomes, the possibility of serious adverse events, and the requirements for long-term surveillance.

Gene therapy illustrates this issue most clearly. The potential for durable benefit may create strong expectations, but the possibility of severe hepatic toxicity and the complexity of post-treatment monitoring necessitate balanced communication. The 2025 Elevidys regulatory actions demonstrate why informed consent should incorporate evolving pharmacovigilance information rather than rely exclusively on the initial approval narrative. Equity is another important ethical dimension. Molecular therapies may be extraordinarily expensive and may require specialized diagnostic testing and treatment infrastructure. If access is determined primarily by geographic or socioeconomic status, the development of precision therapeutics could paradoxically widen disparities in rare-disease care. Ethically robust DMD pharmacotherapy therefore requires transparency, equitable access, respect for patient autonomy, proportionality of risk and benefit, protection of vulnerable pediatric patients, and continuous reassessment of emerging safety information.

Evidence Hierarchy and Limitations of the Current Therapeutic Knowledge Base

The evidence base for DMD pharmacotherapy remains heterogeneous. Some interventions have decades of clinical experience, whereas others have limited post-authorization exposure. Consequently, therapeutic evidence should not be interpreted as a single homogeneous body of knowledge.

Corticosteroids possess extensive clinical experience and multiple lines of functional evidence. Vamorolone has controlled trial evidence supporting its use as a corticosteroid alternative. Givinostat has controlled clinical evidence and regulatory authorization for patients meeting the specified age criterion. Exon-skipping therapies possess molecular and clinical evidence but remain associated with genotype-specific applicability and regulatory requirements for confirmatory evidence. Gene therapy offers a potentially transformative mechanism but has a comparatively complex and evolving long-term safety evidence base. Another limitation concerns endpoint heterogeneity. Studies may measure dystrophin expression, North Star Ambulatory Assessment, timed function tests, muscle strength, pulmonary function, cardiac imaging, or other outcomes. These endpoints are not interchangeable. A therapy may produce measurable molecular activity without demonstrating an equivalent effect on long-term patient-centered outcomes.

The DMD's progressive natural history makes randomized long-term comparisons challenging. Placebo-controlled studies may become ethically and practically difficult as effective therapies become established. Historical controls and natural-history cohorts therefore assume increasing importance, but they introduce methodological uncertainty. The most scientifically defensible interpretation is consequently one that explicitly recognizes differences in evidence maturity rather than assigning equivalent certainty to all therapies.

Clinical Pharmacology of Longitudinal Treatment

A central conclusion from the integrated analysis is that DMD treatment should be regarded as longitudinal pharmacotherapy rather than episodic prescribing. The therapeutic objective is to maintain biological and functional reserve across decades.

This approach requires periodic reassessment of treatment efficacy, tolerability, organ-system status, adherence, growth, body composition, bone health, cardiac function, respiratory function, renal function, hepatic function, and laboratory biomarkers.

For corticosteroids, longitudinal management includes monitoring of growth, weight, blood pressure, glucose metabolism, bone health, cataracts, behavioral effects, and adrenal suppression.

For ASOs, longitudinal management includes infusion tolerance, renal monitoring, treatment adherence, and reassessment of functional trajectory.

For givinostat, platelet counts, triglycerides, gastrointestinal tolerability, and QT-related considerations must remain under surveillance.

For gene therapy, surveillance must be particularly intensive during the period surrounding administration and thereafter, with emphasis on hepatic function and other treatment-specific risks. Current FDA information explicitly identifies hepatic impairment and active hepatic infection as important limitations because of the risk of serious liver injury. This longitudinal perspective also changes the meaning of treatment success. Success should not be defined solely as improvement in muscle strength. Stabilization of function, delayed loss of ambulation, preservation of respiratory reserve, prevention of cardiomyopathy progression, reduction of fractures, maintenance of independence, and improved quality of life may all constitute clinically meaningful therapeutic achievements.

Integration of Pharmacology, Toxicology, and Pharmacovigilance

The most significant conceptual result of this review is that pharmacology, toxicology, and pharmacovigilance should not be treated as separate analytical domains. In DMD, they are functionally interconnected.

Pharmacodynamics explains therapeutic mechanisms; pharmacokinetics determines exposure; toxicology identifies biological harm; pharmacovigilance evaluates rare and delayed adverse events; and clinical pharmacology integrates these dimensions into individualized treatment decisions. This integration is particularly important in pediatric rare disease because the therapeutic window may be influenced by developmental pharmacology, body composition, organ maturation, chronic exposure, and disease progression. Furthermore, DMD patients may receive treatment for many years, meaning that cumulative toxicity can become as clinically important as acute adverse reactions. The contemporary pharmacotherapeutic model should therefore be represented as a continuous cycle: molecular diagnosis → therapeutic selection → dose and exposure optimization → pharmacodynamic response assessment → toxicity monitoring → pharmacovigilance → treatment adaptation. Such a framework is consistent with the broader evolution of DMD care toward multidisciplinary and evidence-based management. Consensus care recommendations emphasize coordinated cardiac, respiratory, bone, orthopaedic, and neuromuscular management rather than isolated treatment of muscle weakness.

Implications for Future Therapeutic Development

The present therapeutic architecture establishes several priorities for future research. First, next-generation molecular therapies should seek broader genetic applicability and greater dystrophin restoration. Second, RNA therapeutics may benefit from improved tissue penetration, increased exon specificity, reduced dosing frequency, and enhanced intracellular delivery. Third, gene therapy research must address immune barriers, hepatotoxicity, vector dose, durability, and the development of safer delivery platforms. Fourth, pharmacological modulation of fibrosis, inflammation, regeneration, mitochondrial dysfunction, and muscle metabolism remains scientifically relevant because dystrophin deficiency activates multiple secondary pathogenic pathways. A successful future therapeutic strategy may therefore require simultaneous or sequential intervention at several mechanistic levels. Fifth, biomarker development is essential. Reliable biomarkers capable of predicting treatment response, toxicity susceptibility, disease progression, and therapeutic durability could substantially improve precision pharmacotherapy.

The future trials should increasingly prioritize patient-centered outcomes. Molecular and histological endpoints remain valuable, but their ultimate significance depends upon whether they translate into preservation of meaningful function, respiratory capacity, cardiac health, independence, and survival.

Clinical Interpretation

Taken together, the results indicate that contemporary DMD pharmacotherapy has progressed from nonspecific disease attenuation toward mechanistically differentiated precision therapeutics. Corticosteroids remain foundational because of their broad clinical effects and extensive evidence base. Vamorolone introduces a more pharmacologically differentiated steroid option. Exon-skipping ASOs establish genotype-dependent RNA therapeutics. Givinostat broadens treatment beyond steroid and mutation-specific approaches through epigenetic modulation. Gene replacement introduces a potentially durable molecular intervention but simultaneously creates unprecedented pharmacosafety challenges.

The major therapeutic advance is therefore not attributable to a single drug. It lies in the expansion of the therapeutic repertoire and the ability to match distinct pharmacological mechanisms to distinct biological and clinical phenotypes.

At the same time, the increasing complexity of the therapeutic armamentarium creates new responsibilities. Greater mechanistic sophistication does not eliminate uncertainty. Molecular precision does not guarantee clinical benefit. Regulatory approval does not eliminate the need for post-marketing surveillance. Novelty does not equate to superiority. High-cost therapy does not automatically constitute high-value therapy. The appropriate therapeutic decision must emerge from systematic evaluation of efficacy, evidence certainty, toxicity, patient characteristics, feasibility, and long-term outcomes.

The Elevidys experience provides a particularly important illustration. The therapy represents a major technological advance, yet subsequent fatal acute liver failure reports in non-ambulatory patients resulted in substantial regulatory revision and a boxed warning. This demonstrates that innovation and pharmacosafety must progress simultaneously.

The therapeutic future of DMD is consequently unlikely to be defined by a single universally dominant intervention. Rather, it will probably involve increasingly individualized combinations of molecular correction, anti-inflammatory and epigenetic modulation, cardiac protection, respiratory support, skeletal preservation, rehabilitation, and precision pharmacovigilance.

The comprehensive analysis establishes that DMD represents one of the most rapidly evolving fields of modern neuromuscular pharmacology. The therapeutic armamentarium now encompasses conventional glucocorticoids, modified corticosteroid pharmacology, sequence-specific antisense therapeutics, epigenetic intervention, and AAV-mediated gene replacement, supplemented by pharmacological management of cardiomyopathy, respiratory complications, skeletal disease, endocrine abnormalities, pain, and other secondary manifestations.

The principal pharmacological determinant of treatment selection is increasingly the relationship between molecular mechanism and individual patient phenotype. Genetic characterization determines eligibility for exon-specific therapies; age and disease stage influence therapeutic timing; ambulatory status has become particularly important for advanced gene therapy; cardiac and respiratory phenotypes determine supportive pharmacotherapy; and cumulative corticosteroid exposure influences endocrine and skeletal risk.

The pharmacodynamic landscape is similarly heterogeneous. Conventional steroids alter receptor-mediated transcription; ASOs modify RNA splicing; givinostat alters epigenetic regulation; and gene therapy introduces a therapeutic transgene. These mechanisms require different pharmacokinetic models, different biomarkers, different monitoring strategies, and different interpretations of therapeutic success.

The toxicological landscape is equally differentiated. Chronic corticosteroid therapy is characterized primarily by cumulative systemic toxicity; ASOs require attention to infusion and renal safety; givinostat requires laboratory and cardiac monitoring; and gene therapy introduces potentially severe immune-mediated and hepatic risks. The appropriate response is therefore not

avoidance of pharmacological risk but systematic risk characterization, mitigation, monitoring, and transparent communication.

Pharmacokinetics and Safety

ASOs exhibit distinctive pharmacokinetic properties: large molecular size, charge characteristics, protein interactions, tissue distribution, intracellular trafficking, and renal elimination. Metabolism occurs through nuclease-mediated pathways rather than classical hepatic CYP450 biotransformation. Repeated intravenous administration is required, and renal handling is particularly relevant to safety surveillance.

Principal safety concerns include infusion-related reactions, renal toxicity requiring monitoring, treatment burden from repeated administration, and uncertainty surrounding the magnitude of long-term clinical benefit. (See Table-3).

Table 3. Exon-Skipping Antisense Oligonucleotide Therapies

Agent	Target Exon	Eligible Population	Approval	Key Evidence	Major Limitations
Eteplirsen	Exon 51	~13% of DMD	FDA 2016 (AA)	Dystrophin restoration	Clinical benefit uncertainty
Golodirsen	Exon 53	~8% of DMD	FDA 2019 (AA)	Dystrophin restoration	Confirmatory evidence required
Viltolarsen	Exon 53	~8% of DMD	FDA 2020 (AA)	Dystrophin restoration	Long-term data evolving
Casimersen	Exon 45	~8% of DMD	FDA 2021 (AA)	Dystrophin restoration	Very restricted population

Abbreviation: AA, accelerated approval.

The therapeutic objective should consequently be formulated as preservation of lifelong biological and functional capacity rather than short-term improvement in muscle performance alone. This requires integrating molecular pharmacology with clinical pharmacology, pharmacokinetics, pharmacodynamics, toxicology, pharmacovigilance, cardiology, pulmonology, endocrinology, orthopaedics, rehabilitation, genetics, and ethical medicine.

The contemporary pharmacotherapeutic armamentarium of DMD should be understood as an evolving precision-treatment ecosystem rather than a static list of medications. Its continued advancement will depend upon increasingly accurate molecular stratification, robust comparative effectiveness research, harmonized clinical endpoints, longitudinal real-world evidence, improved biomarkers, vigilant post-marketing surveillance, and ethically responsible incorporation of emerging technologies. The ultimate criterion of therapeutic success remains the capacity to preserve meaningful function, prolong survival, minimize treatment-associated harm, and improve quality of life while respecting the biological individuality and autonomy of each person living with DMD.

ANALYSIS

Analytical Framework and Scope of theAnalytic Synthesis

The present meta-analytic synthesis was designed to evaluate the comparative and complementary evidence surrounding the contemporary pharmacotherapeutic management of Duchenne muscular dystrophy (DMD), with particular emphasis on therapeutic efficacy, functional preservation, safety, and pharmacological heterogeneity. Because DMD is a rare, genetically determined, progressive neuromuscular disorder and because currently available interventions operate through fundamentally different biological mechanisms, a single global pooled effect estimate across all therapeutic modalities would be methodologically inappropriate. Instead, the evidence was interpreted according to pharmacological class, molecular target, therapeutic mechanism, clinical endpoint, comparator, and evidentiary maturity.

The analytical hierarchy comprised five major pharmacotherapeutic domains: conventional glucocorticoids; modified corticosteroid therapy with vamorolone; exon-skipping antisense oligonucleotides; epigenetic therapy with givinostat; and AAV-mediated micro-dystrophin gene replacement with delandistrogene moxeparvovec. Supportive cardiac, respiratory, skeletal, endocrine, and symptomatic pharmacotherapy was incorporated into the interpretative framework but was not artificially pooled with disease-modifying interventions because these therapies address different clinical objectives and lack sufficiently homogeneous randomized datasets.

The principal efficacy outcomes considered were changes in 6-minute walk distance (6MWD), North Star Ambulatory Assessment (NSAA), timed rise from the floor or supine position, timed stair climbing, timed walking/running, pulmonary function, preservation of ambulation, and molecular measures such as dystrophin or micro-dystrophin expression. Safety outcomes included serious adverse drug reactions, weight gain, endocrine toxicity, behavioral effects, cataracts, renal toxicity, thrombocytopenia, gastrointestinal toxicity, hepatic injury, immune-mediated adverse events, and treatment discontinuation.

The methodological interpretation followed a hierarchy in which randomized controlled trials and prospectively planned meta-analyses were accorded greater evidentiary weight than post hoc analyses, uncontrolled longitudinal studies, and historical comparisons. This distinction is particularly important in DMD because natural-history trajectories are strongly age-dependent and because relatively small sample sizes may generate unstable estimates of treatment effects.

Quantitative Evidence for Corticosteroid Therapy

The strongest historical quantitative evidence in DMD concerns corticosteroid therapy. Earlier systematic reviews and meta-analyses demonstrated that prednisone/prednisolone significantly improves muscle strength and motor function over approximately six months compared with placebo. Improvements were reported across measures including time to rise from the floor, timed walking, stair climbing, weight-lifting capacity, leg function, and forced vital capacity. The evidence supported approximately 0.75 mg/kg/day prednisone or prednisolone as a clinically effective regimen.

The significance of these findings extends beyond statistical significance. DMD is characterized by progressive loss of muscle reserve, and even relatively modest preservation of functional capacity may translate into delayed disability when treatment is sustained over years. The therapeutic effect of glucocorticoids is therefore better interpreted as attenuation of disease progression rather than restoration of normal neuromuscular physiology.

A particularly informative quantitative comparison concerns deflazacort versus prednisone/prednisolone. A pooled analysis of placebo-arm participants from two multicenter trials included 231 ambulatory boys with DMD receiving stable corticosteroid regimens. Over 48

weeks, deflazacort-treated participants experienced a significantly smaller decline in 6MWD, with a pooled adjusted difference of 28.3 m in favor of deflazacort. The corresponding differences were approximately 2.9 seconds for rise from supine, 2.3 seconds for four-stair climbing, and 2.9 points for the linearized NSAA score.

A separate meta-analysis combining placebo data from two clinical trials similarly reported a 39.54-m least-squares mean difference in 6MWD favoring deflazacort over prednisone/prednisolone, together with clinically meaningful improvements in timed stair-climbing measures. These findings collectively suggest that deflazacort may provide somewhat greater preservation of ambulatory function than prednisone/prednisolone under the analyzed conditions.

However, these estimates should not be interpreted as definitive proof that deflazacort is universally superior. The available comparisons have important methodological limitations, including retrospective or post hoc analyses, differences in patient populations, corticosteroid exposure history, baseline characteristics, and trial-specific designs. One earlier randomized comparison found broadly similar motor efficacy but substantially greater weight gain with prednisone than deflazacort.

Consequently, the meta-analytic evidence supports a more nuanced conclusion: both prednisone/prednisolone and deflazacort are effective glucocorticoid therapies, while deflazacort may offer a favorable functional and anthropometric profile in selected patients.

Quantitative Comparison of Prednisone and Deflazacort

The comparative corticosteroid evidence illustrates a fundamental principle of clinical pharmacology: efficacy and tolerability cannot be separated when assessing chronic therapy. The American Academy of Neurology evidence synthesis concluded that prednisone should be offered to improve strength and pulmonary function and that deflazacort may be offered for strength, timed motor function, and delay of loss of ambulation. The guideline estimated a potential delay in loss of ambulation of approximately 1.4–2.5 years with deflazacort, while acknowledging limitations in the evidence base.

The comparative evidence therefore suggests that the principal differences between conventional corticosteroids are not necessarily categorical efficacy differences but rather differences in adverse-effect burden and individual patient response.

Prednisone is associated with a comparatively greater tendency toward weight gain. In DMD, this is clinically consequential because excess body mass increases the mechanical workload imposed on weakened skeletal muscle and may accelerate functional deterioration. Deflazacort may therefore confer an indirect functional advantage through reduced treatment-associated weight gain.

Conversely, deflazacort has been associated with greater cataract risk. This illustrates why drug selection cannot be based exclusively on motor endpoints. A treatment that preserves several meters of walking distance but produces clinically important ophthalmological or endocrine toxicity may not represent the optimal individualized therapy. The meta-analytic evidence consequently favors individualized corticosteroid selection rather than a universal hierarchy. (See Table-4).

Table 4. Quantitative Evidence for Major Corticosteroid Comparisons

Therapeutic comparison	Population/evidence	Principal outcome	Quantitative finding	Interpretation
Prednisone/prednisolone vs placebo	Multiple RCTs/meta-analysis	Strength and motor function	Significant improvement over approximately 6 months	Strongest historical pharmacological evidence in DMD
Prednisone 0.75 mg/kg/day	RCT evidence	Strength/function	Superior to lower-dose regimens on most assessed outcomes	Supports conventional daily dosing
Deflazacort vs prednisone/prednisolone	231 participants from multicenter trial datasets	6MWD	+28.3 m adjusted pooled difference favoring deflazacort	Suggests reduced functional decline
Deflazacort vs prednisone/prednisolone	Retrospective pooled analysis	6MWD	+39.54 m least-squares mean difference	Supports possible functional advantage
Deflazacort vs prednisone	Randomized comparative evidence	Body weight	Lower weight gain with deflazacort	Important long-term tolerability advantage
Corticosteroids overall	Systematic reviews	Strength, timed function, FVC	Consistent short-term functional benefit	Supports corticosteroids as foundational therapy

The quantitative evidence summarized above is derived from published meta-analyses and randomized or comparative clinical studies rather than from a newly calculated pooled dataset.

Meta-Analytic Evidence for Vamorolone

Vamorolone represents a pharmacologically differentiated corticosteroid strategy designed to preserve anti-inflammatory/glucocorticoid activity while potentially reducing selected chronic steroid toxicities. Its quantitative evidence base is considerably newer than that of prednisone and deflazacort.

A randomized, double-blind, placebo- and prednisone-controlled clinical trial evaluated boys aged 4 to <7 years with genetically confirmed DMD who had not previously received corticosteroids. Participants were randomized to placebo, prednisone 0.75 mg/kg/day, or vamorolone at 2 or 6 mg/kg/day. The trial demonstrated preservation of motor efficacy with vamorolone relative to placebo, while providing evidence supporting a potentially differentiated safety profile.

A meta-analysis incorporating four studies and 277 participants reported significant improvements in several timed motor measures with vamorolone compared with placebo. Specifically, pooled differences favored vamorolone for time-to-stand velocity, time-to-run/walk velocity, and time-to-climb velocity. The same analysis identified differences in selected outcomes when vamorolone was compared with conventional glucocorticoids.

A subsequent systematic review and meta-analysis including prospective studies reported significant effects of vamorolone on timed functional outcomes, although the quantitative synthesis was based on only two study reports involving 210 participants. At 2 mg/kg/day, statistically significant improvements were identified in timed four-stair climbing and standing-from-supine velocity.

A further network meta-analysis examined different vamorolone doses and reported statistically significant differences favoring 6 mg/kg/day over 2 mg/kg/day for several motor endpoints, including 6MWT, while indirect comparisons did not demonstrate clear differences in BMI z-score among vamorolone and conventional glucocorticoid regimens.

The quantitative evidence therefore supports vamorolone as an effective corticosteroid-class therapy but does not justify describing it as categorically superior to prednisone or deflazacort across all clinically meaningful endpoints. Cross-trial comparisons are inherently vulnerable to confounding from differences in age, baseline motor function, trial duration, outcome ascertainment, steroid exposure, and statistical methodology.

Comparative Meta-Analytic Interpretation of Conventional and Modified Corticosteroids

When the available evidence is considered collectively, three conclusions emerge:

First, corticosteroid treatment as a class has a robust functional signal. This conclusion is supported by multiple randomized trials and systematic reviews.

Second, the magnitude of benefit differs among functional outcomes. Strength and timed motor measures appear more consistently responsive than long-term endpoints such as survival or permanent preservation of ambulation, for which randomized evidence is more difficult to obtain.

Third, safety and tolerability may provide clinically important differentiation between steroid preparations. Prednisone has greater association with weight gain, whereas deflazacort has greater cataract concern. Vamorolone has been developed to modify this efficacy–toxicity relationship and appears to preserve glucocorticoid-associated motor benefit while potentially altering selected adverse-effect profiles. The evidence supports a therapeutic paradigm in which the choice of corticosteroid is individualized according to expected motor benefit, body composition, growth, skeletal health, behavior, ophthalmological status, metabolic risk, and treatment adherence.

Analysis of Exon-Skipping Antisense Oligonucleotide Therapy

The quantitative evidence for exon-skipping therapy is substantially more complex. An early systematic review and meta-analysis of double-blind randomized trials included five studies involving 322 participants and evaluated exon-skipping agents including eteplirsen and drisapersen. The pooled analysis did not demonstrate a statistically significant improvement in 6MWD at 24 weeks, with a mean difference of −9.16 m (95% CI −21.94 to 3.62), nor a significant improvement in NSAA, with a mean difference of 1.20 points (95% CI −2.35 to 4.75).

These findings should not be interpreted as evidence that all exon-skipping approaches are ineffective. The analyzed studies included heterogeneous compounds, mutation-specific populations, dosing regimens, and methodological characteristics. The analysis also predates several later exon-skipping developments and regulatory decisions.

The fundamental limitation of conventional meta-analysis in this therapeutic area is that the molecular eligibility criteria are highly restrictive. Eteplirsen, golodirsen, viltolarsen, and

casimersen are not interchangeable drugs administered to a common DMD population. They target different exons and therefore different subsets of the DMD genotype spectrum.
The most recent literature emphasizes that the four FDA-approved exon-skipping ASOs represent a pioneering but still evolving therapeutic framework characterized by restricted patient eligibility, challenges in extrapolating molecular surrogate responses to clinical outcomes, and substantial treatment burden. This distinction is crucial for meta-analysis. Pooling all ASOs into one estimate may produce statistical simplicity at the expense of biological validity. The appropriate unit of interpretation is therefore the individual molecular intervention and its genotype-defined population. (See Table-5)

Table 5. Quantitative Evidence for Molecular Disease-Modifying Therapies

Therapy/class	Molecular strategy	Principal quantitative evidence	Major limitation
Eteplirsen	Exon 51 skipping	Early pooled RCT evidence did not show significant 24-week 6MWD improvement	Small trials, molecularly restricted population
Golodirsen	Exon 53 skipping	Evidence centered on dystrophin restoration and functional follow-up	Restricted genotype and confirmatory-evidence requirements
Viltolarsen	Exon 53 skipping	Molecular dystrophin restoration with ongoing clinical-evidence development	Genotype-specific eligibility
Casimersen	Exon 45 skipping	Dystrophin restoration and clinical development data	Very limited eligible population
Givinostat	HDAC inhibition	Phase 3 randomized evidence supports slowing selected functional decline	Laboratory, gastrointestinal, platelet and cardiac monitoring
Delandistrogene moxeparvovec	AAV micro-dystrophin gene transfer	Primary EMBARK endpoint not statistically significant at 52 weeks; secondary timed outcomes favored treatment	Long-term efficacy and safety uncertainty; substantial biological risk

The ASO evidence requires particularly cautious interpretation because molecular target engagement and dystrophin restoration are not equivalent to demonstrated preservation of clinically meaningful function.

Givinostat: Quantitative Evidence for Epigenetic Pharmacotherapy

Givinostat provides one of the strongest examples of a nonsteroidal pharmacological approach with randomized phase 3 evidence. The EPIDYS study was a multicenter, randomized, double-blind, placebo-controlled phase 3 trial conducted at 41 tertiary centers across 11 countries. The study evaluated the efficacy and safety of givinostat in boys with DMD receiving background corticosteroid therapy.
The pharmacological significance of this study extends beyond its numerical endpoint estimates. Givinostat does not require a specific exon-level genetic mutation for treatment eligibility and therefore addresses a broader molecular population than exon-specific ASOs. Its mechanism is based on histone deacetylase inhibition and modulation of transcriptional and inflammatory pathways.

The quantitative evidence indicates that givinostat can slow aspects of functional deterioration when administered alongside background corticosteroid therapy. This is conceptually different from replacing corticosteroids. Rather, it supports a layered pharmacotherapeutic model in which a nonsteroidal disease-modifying agent is added to established steroid treatment.

Safety is an equally important component of the meta-analytic interpretation. Givinostat therapy requires attention to platelet counts, triglycerides, gastrointestinal adverse effects, and electrocardiographic risk. A 2026 meta-analysis of pharmacological treatment safety identified heterogeneity in adverse-drug-reaction incidence and reported dose- and duration-dependent safety patterns for corticosteroids, while also identifying differences in selected adverse events between corticosteroid regimens and givinostat-containing treatment. Thus, the therapeutic value of givinostat should be interpreted through a benefit–risk framework rather than through functional efficacy alone.

Meta-Analytic Evidence for Delandistrogene Moxeparvovec

Gene replacement therapy presents the greatest methodological challenge for meta-analysis. Delandistrogene moxeparvovec is administered as a single intravenous gene-transfer intervention, and its biological effect depends on vector biodistribution, muscle transduction, micro-dystrophin expression, immune response, and persistence of expression. These properties make direct comparison with conventional daily pharmacotherapy inherently problematic.

The phase 3 EMBARK randomized trial enrolled ambulatory boys aged 4 to <8 years. At week 52, the primary endpoint—change in NSAA—did not reach statistical significance. The least-squares mean change was 2.57 points with delandistrogene moxeparvovec versus 1.92 points with placebo, yielding a between-group difference of 0.65 points (95% CI −0.45 to 1.74; P=0.2441). Nevertheless, several secondary timed functional endpoints numerically favored treatment, and micro-dystrophin expression was demonstrated. This distinction is critical. A meta-analysis must not transform secondary-endpoint signals into proof of primary-endpoint efficacy. The appropriate conclusion from EMBARK is that the study did not meet its prespecified primary functional endpoint at one year, while selected secondary outcomes suggested possible slowing of disease progression. More recent two-year follow-up reported sustained stabilization or slowing of disease progression in comparison with a matched external-control cohort and persistence of micro-dystrophin expression, without new safety signals between weeks 52 and 104. However, this comparison is not equivalent to a contemporaneous randomized placebo-controlled analysis and therefore carries greater potential for residual confounding.

A 2025 systematic review and meta-analysis involving five trials and 190 participants reported improvement in several motor outcomes at one year, including a 2.63-point improvement in NSAA in pre–post analyses and a −0.64-second difference in TTSTAND compared with placebo. However, severe adverse events occurred in approximately 10–15% of participants, and the authors emphasized the need for longer-term evidence. A newer 2026 meta-analysis including 302 participants from four studies, including two randomized trials, reported an approximately 2.48-point improvement in NSAA, a −0.85-second improvement in time to rise, and an increase in muscle dystrophin content. Nevertheless, the authors specifically noted substantial heterogeneity.

The apparent discrepancy between the pivotal randomized trial and subsequent pooled analyses illustrates the importance of study design. Pre–post analyses are susceptible to regression to the mean and natural-history effects, while external-control analyses may be affected by differences in baseline prognosis. Consequently, the strongest causal inference remains the randomized comparison.

Safety Meta-Analysis and Pharmacosafety Interpretation

Safety is not a secondary endpoint in DMD pharmacotherapy. Given the prolonged duration of treatment and the pediatric age at initiation, cumulative toxicity may substantially influence the overall clinical value of an intervention.

A recent meta-analysis of clinical trials evaluating pharmacological treatments for DMD included 13 trials and identified significant heterogeneity in adverse-drug-reaction incidence. Prednisone and prednisolone showed comparatively high proportions of serious adverse reactions, while dose- and duration-dependent adverse-event patterns were identified for deflazacort and prednisone. The analysis also identified differences in selected gastrointestinal and hematological adverse events between corticosteroid monotherapy and givinostat-containing treatment. These findings support an important pharmacological proposition: safety cannot be reduced to a binary classification of “safe” versus “unsafe.” The toxicity of a treatment is multidimensional and dependent upon exposure duration, dose, age, baseline susceptibility, concomitant medication, and monitoring intensity.

Corticosteroid toxicity is predominantly cumulative and systemic. ASO toxicity is more closely related to administration burden, infusion reactions, renal considerations, and repeated exposure. Givinostat introduces hematological, gastrointestinal, lipid, and electrophysiological monitoring requirements. Gene therapy introduces immune-mediated and hepatic risks that may be acute and potentially severe.

The pharmacosafety profile therefore influences therapeutic selection at the same level as efficacy. (See Table-6).

Table 6. Integrated Efficacy–Safety Interpretation

Therapeutic modality	Efficacy signal	Major safety domains	Evidence maturity	Interpretation
Prednisone/prednisolone	Robust improvement in strength and motor function	Weight gain, growth suppression, behavioral effects, osteoporosis, adrenal suppression	High historical experience	Foundational therapy
Deflazacort	Functional preservation; possible advantage over prednisone	Cataracts, steroid toxicity, adrenal suppression	Moderate–high	Important alternative corticosteroid
Vamorolone	Motor efficacy versus placebo; potentially differentiated toxicity	Steroid-associated endocrine/metabolic effects remain relevant	Moderate and expanding	Promising modified corticosteroid

Exon-skipping ASOs	Molecular dystrophin restoration; clinical benefit varies by agent	Infusion reactions, renal considerations, treatment burden	Heterogeneous	Precision therapy for genotype-defined populations
Givinostat	Functional slowing demonstrated in phase 3	Thrombocytopenia, diarrhea, hypertriglyceridemia, QTc concerns	Moderate–high	Important genotype-independent nonsteroidal option
Delandistrogene moxeparvovec	Micro-dystrophin expression; selected secondary functional benefits	Hepatic injury, immune-mediated events, thrombocytopenia, myocarditis/myositis	Rapidly evolving	Potentially durable but high-complexity therapy

Safety and Pharmacovigilance

The evolution of the Elevidys safety profile is one of the most important pharmacovigilance findings in contemporary DMD therapy. In 2025, the FDA reported fatal acute liver failure in non-ambulatory pediatric patients following treatment and subsequently required a boxed warning and revised the indication to restrict use to ambulatory patients aged four years and older. This development illustrates the dynamic nature of pharmacosafety knowledge in advanced therapies: pre-approval clinical trials may not fully characterize rare, severe, or phenotype-dependent adverse events.

Key safety considerations include immune activation, hepatic injury, thrombocytopenia, and potentially severe immune-mediated adverse events. Patient selection, baseline serological assessment, corticosteroid immunomodulation, hepatic monitoring, and long-term pharmacovigilance constitute essential elements of responsible gene-therapy deployment.

Table 7. Comparative Efficacy–Safety Profile of Disease-Modifying Therapies

Therapeutic Modality	Efficacy Signal	Evidence Maturity	Major Safety Domains	Clinical Interpretation
Prednisone/Prednisolone	Robust motor improvement	High historical	Weight gain, growth suppression, behavioral effects, osteoporosis, adrenal suppression	Foundational therapy
Deflazacort	Functional preservation; possible advantage	Moderate-High	Cataracts, steroid toxicity, adrenal suppression	Important alternative
Vamorolone	Motor efficacy; differentiated toxicity	Moderate-expanding	Steroid-associated effects remain relevant	Promising modified corticosteroid
Exon-skipping ASOs	Molecular restoration; variable clinical benefit	Heterogeneous	Infusion reactions, renal considerations, treatment burden	Precision therapy for genotype-defined populations
Givinostat	Functional slowing in phase 3	Moderate-High	Thrombocytopenia, diarrhea, hypertriglyceridemia, QTc	Genotype-independent nonsteroidal option
Gene Therapy	Micro-dystrophin expression; secondary benefits	Rapidly evolving	Hepatic injury, immune-mediated events, thrombocytopenia	Potentially durable but high-complexity

Heterogeneity as a Central Statistical and Biological Problem

Heterogeneity is one of the most important limitations affecting quantitative synthesis in DMD. It arises from at least six major sources: age, baseline motor function, genotype, corticosteroid exposure, outcome selection, and study design.

Age is particularly important because motor decline accelerates at different stages of the disease. A treatment effect observed in four-year-old ambulatory children cannot automatically be extrapolated to adolescents with advanced weakness.

Genotype creates another layer of heterogeneity. Exon-skipping therapies are inherently restricted to molecularly defined populations, while givinostat and corticosteroids have substantially broader applicability.

Corticosteroid exposure is another major confounder. Many modern trials are conducted in patients already receiving stable steroids, meaning that the incremental treatment effect represents add-on efficacy rather than absolute efficacy versus untreated disease.

Outcome heterogeneity also complicates quantitative synthesis. NSAA, 6MWD, TTSTAND, TTCLIMB, 10MWR, pulmonary function, dystrophin expression, and loss of ambulation measure different dimensions of disease. A statistically significant result in one endpoint should not automatically be interpreted as equivalent to a statistically significant result in another. The study-design heterogeneity is substantial. Randomized controlled trials provide stronger causal inference than open-label studies, natural-history comparisons, or external-control analyses.

Publication Bias and Small-Study Effects

The relatively small size of DMD trials creates an increased risk of small-study effects. Rare-disease therapeutic development often involves highly selected populations and specialized centers, which may limit external validity.

Publication bias is also a potential concern because studies with favorable molecular or clinical findings may be more likely to reach publication. Conversely, negative or inconclusive trials may be underrepresented in the published literature.

The early exon-skipping literature illustrates this problem particularly clearly. The systematic review found only a small number of randomized trials and identified methodological concerns, including high risk of bias in some studies. Consequently, pooled estimates should be interpreted with appropriate skepticism. A narrow confidence interval does not necessarily eliminate uncertainty if the underlying studies are small, heterogeneous, or methodologically limited.

Surrogate Endpoints and Clinical Meaningfulness

A recurring theme across the meta-analytic evidence is the distinction between biological efficacy and clinical effectiveness. Dystrophin or micro-dystrophin expression is a biologically compelling endpoint because dystrophin deficiency is the fundamental molecular defect in DMD. However, the magnitude of protein restoration necessary to produce clinically meaningful benefit remains incompletely established for every therapeutic platform. Similarly, improved RNA exon skipping confirms molecular target engagement but does not automatically establish preservation of ambulation or survival. This distinction is particularly important for regulatory interpretation. Molecular surrogate endpoints may support accelerated development, but confirmatory clinical evidence remains essential. The continued evolution of the evidence base for exon-skipping agents demonstrates why long-term functional studies and real-world evidence remain necessary. The most clinically meaningful endpoints are ultimately those reflecting preservation of independence, ambulation, respiratory function, cardiac health, quality of life, and survival.

Comparative Therapeutic Effectiveness

A direct network meta-analysis across all contemporary DMD treatments would currently be difficult to justify scientifically because the therapies are not evaluated within a common evidence architecture.

Prednisone, deflazacort, and vamorolone can reasonably be compared within a corticosteroid framework. Givinostat can be evaluated against placebo plus background corticosteroid

treatment. Exon-skipping ASOs require mutation-specific interpretation. Gene therapy requires evaluation within a pediatric ambulatory population and consideration of both randomized and external-control evidence.

Therefore, the evidence supports a hierarchical rather than linear ranking:

- Corticosteroids have the longest-established evidence for broad functional benefit.
- Vamorolone provides a modified corticosteroid strategy with increasing evidence.
- Givinostat provides a nonsteroidal, genotype-independent pharmacological mechanism with phase 3 evidence.
- Exon-skipping ASOs provide highly precise molecular treatment for selected genotypes, but their clinical evidence remains heterogeneous.
- Gene replacement therapy offers potentially durable molecular correction but currently carries substantial uncertainty concerning long-term effectiveness and safety.

This hierarchy does not imply that one therapy is clinically superior in every patient. Rather, it describes relative evidence maturity and biological specificity.

Clinical and Pharmacological Implications of the Meta-Analytic Findings

The integrated findings support a transition from the concept of "the DMD drug" toward an individualized therapeutic architecture. The appropriate therapeutic strategy is determined by the intersection of genetic eligibility, disease stage, ambulatory status, previous treatment, organ-system involvement, anticipated benefit, adverse-event susceptibility, and patient preferences.

For corticosteroids, the meta-analysis supports continued use as foundational therapy while emphasizing individualized selection between prednisone, deflazacort, and vamorolone.

For ASOs, genetic confirmation should precede therapeutic selection. Exon-skipping treatment is inherently a genotype-guided intervention and should not be prescribed solely on the basis of clinical DMD diagnosis.

For givinostat, the evidence supports its role as a nonsteroidal disease-modifying therapy for eligible patients, but its laboratory monitoring requirements must be integrated into routine care.

For gene therapy, the evidence supports cautious use in appropriately selected ambulatory patients while recognizing the limitations of the pivotal randomized evidence and the continuing evolution of post-marketing safety information.

Pharmacovigilance and Long-Term Evidence Generation

The meta-analysis demonstrates that clinical trials alone cannot resolve all uncertainties associated with DMD pharmacotherapy. Long-term registries and real-world evidence are therefore essential.

This is especially true for gene therapy, where rare hepatic or immune-mediated complications may become apparent only after broader population exposure. The subsequent evolution of the safety profile of delandistrogene moxeparvovec demonstrates the importance of post-authorization surveillance.

The 2025 regulatory experience with Elevidys, including reports of fatal acute liver failure in non-ambulatory patients and subsequent regulatory restriction, provides a particularly important pharmacovigilance lesson. Advanced therapies require continuous reassessment of benefit and risk as new patient populations are treated. The therapeutic evidence base should therefore be regarded as dynamic rather than finalized at the moment of regulatory approval.

Methodological Limitations of the Present Meta-Analytic Interpretation

Several limitations should be explicitly acknowledged: First, the present synthesis integrates published quantitative analyses rather than reconstructing individual-patient-level datasets. Consequently, it should not be interpreted as a de novo individual-participant meta-analysis. Second, therapeutic modalities were intentionally not combined into one overall pooled effect

because such pooling would violate fundamental assumptions of clinical and biological comparability. Third, several published estimates originate from post hoc analyses, external-control comparisons, or pre–post studies. Such designs have lower causal certainty than randomized placebo-controlled trials. Fourth, DMD trials frequently involve small samples, which limits statistical power for detecting rare adverse events. Fifth, follow-up duration varies substantially. Short-term motor outcomes cannot reliably predict lifelong preservation of ambulation, cardiac function, respiratory independence, or survival. Sixth, publication dates and regulatory status are dynamic. Particularly for gene therapy, emerging post-marketing evidence can materially change the interpretation of benefit–risk balance. Seventh, cross-trial comparisons are susceptible to differences in corticosteroid background therapy, baseline disease severity, age, endpoint definitions, and analytic methodology. These limitations reinforce the necessity of cautious interpretation and prevent overstatement of therapeutic superiority.

Integrated Evidence Table

Table 8. Comprehensive Meta-Analytic Evidence Matrix for Contemporary DMD Pharmacotherapy

Domain	Intervention	Mechanism	Main Quantitative Evidence	Principal benefit	Principal limitation
Corticosteroid	Prednisone/prednisolone	Glucocorticoid receptor modulation	Significant strength/function benefit versus placebo	Broad applicability and extensive evidence	Cumulative systemic toxicity
Corticosteroid	Deflazacort	Glucocorticoid receptor modulation	Meta-analytic 6MWD advantage vs prednisone in selected datasets	Functional preservation; less weight gain	Cataract and steroid-associated toxicity
Modified corticosteroid	Vamorolone	Differentiated glucocorticoid receptor pharmacology	Significant improvement in selected timed motor outcomes	Potentially improved tolerability profile	Long-term comparative evidence remains developing
RNA therapy	Eteplirsen	Exon 51 skipping	Early pooled RCT evidence inconclusive for 6MWD	Genotype-specific dystrophin restoration	Restricted eligibility; clinical-effect uncertainty

RNA therapy	Golodirsen	Exon 53 skipping	Molecular and clinical-development evidence	Precision molecular therapy	Restricted genotype
RNA therapy	Viltolarsen	Exon 53 skipping	Molecular dystrophin restoration	Precision therapy	Restricted genotype and long-term evidence
RNA therapy	Casimersen	Exon 45 skipping	Molecular dystrophin restoration	Genotype-specific intervention	Very restricted eligible population
Epigenetic	Givinostat	HDAC inhibition	Phase 3 randomized evidence	Genotype-independent nonsteroidal option	Platelet, lipid, GI and QT monitoring
Gene therapy	Delandistrogene moxeparvovec	AAVrh74 micro-dystrophin transfer	EMBARK primary endpoint negative at 52 weeks; secondary endpoints favorable	Potentially durable molecular intervention	Hepatic/immune toxicity and efficacy uncertainty
Cardiac support	ACEi/ARB/MRA/beta-blocker	Neurohormonal blockade	Strong mechanistic and consensus evidence	Cardiomyopathy prevention/management	Requires organ-specific monitoring
Bone support	Vitamin D/calcium/bisphosphonates	Bone mineral preservation	Supportive evidence	Reduction of skeletal complications	Does not directly modify dystrophin pathology
Respiratory support	Bronchodilator/mucolytic therapy when indicated	Airway pharmacology	Indication-dependent	Management of concomitant airway disease	Does not reverse respiratory muscle weakness

Analytic Interpretation

The integrated meta-analytic evidence demonstrates that DMD pharmacotherapy has entered a highly differentiated therapeutic era in which clinical efficacy, molecular precision, pharmacokinetic complexity, and safety surveillance must be interpreted simultaneously. The most robust quantitative evidence remains associated with corticosteroids. Prednisone/prednisolone consistently improve strength and motor function, while comparative

analyses suggest that deflazacort may reduce functional decline and weight gain relative to prednisone in selected populations.

Vamorolone has emerged as an important pharmacological evolution of steroid therapy. Meta-analytic evidence supports improvements in several timed motor outcomes, although its comparative superiority across all clinical endpoints remains unproven. The evidence for exon-skipping ASOs is more heterogeneous. The mechanistic rationale is exceptionally strong, but the clinical evidence is constrained by small samples, mutation-specific populations, treatment burden, and uncertainty concerning the magnitude of long-term patient-centered benefit. Early pooled randomized evidence did not demonstrate significant improvement in 6MWD or NSAA at 24 weeks, emphasizing the distinction between molecular plausibility and demonstrated clinical effectiveness.

Givinostat occupies an increasingly important position because it provides a nonsteroidal pharmacological mechanism applicable across genetic variants. Its phase 3 evidence supports a measurable functional benefit, but treatment requires active laboratory and cardiac safety surveillance.

Gene therapy represents the most transformative but also the most uncertain therapeutic domain. Delandistrogene moxeparvovec clearly produces micro-dystrophin expression and has demonstrated favorable signals in selected secondary motor outcomes; however, the pivotal EMBARK study did not achieve its primary NSAA endpoint at 52 weeks. Subsequent two-year analyses and meta-analyses have reported encouraging stabilization or improvement in several functional measures, but these findings remain subject to limitations associated with external controls, heterogeneity, and limited long-term randomized evidence.

The evidence does not support a simplistic ranking of contemporary DMD therapies. Instead, the findings support an individualized pharmacotherapeutic model in which therapeutic choice is determined by the interaction of genotype, age, disease stage, ambulatory status, baseline function, cardiac and respiratory phenotype, prior corticosteroid exposure, safety profile, and evidentiary certainty. The most scientifically defensible conclusion is therefore that contemporary DMD pharmacotherapy has achieved substantial therapeutic diversification without yet resolving the fundamental challenge of lifelong disease modification. Corticosteroids remain the pharmacological foundation; modified corticosteroids seek to optimize the efficacy–toxicity relationship; exon-skipping ASOs introduce genotype-specific RNA therapeutics; givinostat expands nonsteroidal epigenetic treatment; and gene replacement introduces the possibility of durable molecular intervention. However, each modality possesses distinct limitations that prevent substitution of one therapeutic paradigm for another.

The future of DMD treatment will consequently depend less on identification of a single universally superior drug than on rational integration of complementary pharmacological mechanisms. The emerging therapeutic objective is increasingly one of precision longitudinal disease management**,** in which molecular correction, inflammatory modulation, epigenetic regulation, cardioprotection, respiratory preservation, skeletal protection, rehabilitation, and pharmacovigilance are integrated into an individualized therapeutic continuum. From an evidence-based medicine perspective, the central finding is equally important: statistically significant molecular or functional effects must not be interpreted independently of clinical relevance, safety, durability, and methodological certainty. The appropriate interpretation of the contemporary evidence is therefore neither therapeutic pessimism nor uncritical enthusiasm, but a disciplined benefit–risk framework that recognizes genuine advances while explicitly preserving scientific uncertainty where the available data remain incomplete.

The meta-analytic evidence supports a transition from conventional disease attenuation toward increasingly mechanistically targeted DMD pharmacotherapy. Nevertheless, the ultimate standard of therapeutic success remains preservation of meaningful human function: delayed loss of

ambulation, sustained respiratory and cardiac capacity, reduced treatment-associated morbidity, prolonged independence, improved quality of life, and ultimately increased survival. Achieving these objectives will require continued integration of clinical trials, quantitative evidence synthesis, molecular biomarkers, pharmacokinetic–pharmacodynamic modeling, real-world evidence, and rigorous pharmacovigilance throughout the lifespan of individuals living with Duchenne muscular dystrophy.

CONCLUSION

- Duchenne muscular dystrophy (DMD) constitutes a prototypical disorder at the intersection of molecular genetics, neuromuscular pathobiology, translational pharmacology, and precision therapeutics, representing a particularly demanding paradigm for contemporary drug development. The absence or profound deficiency of functional dystrophin initiates a complex and progressively self-reinforcing cascade of sarcolemmal instability, contraction-induced myofiber injury, necrotic and regenerative imbalance, chronic inflammation, mitochondrial dysfunction, aberrant calcium handling, fibroadipose replacement, and ultimately irreversible loss of skeletal- and cardiac-muscle integrity. Accordingly, the pharmacological management of DMD cannot be reduced to correction of a single molecular lesion or suppression of an isolated pathological pathway. Rather, it requires an integrated therapeutic architecture capable of simultaneously addressing the primary genetic defect, downstream molecular pathology, progressive tissue remodeling, multisystem complications, treatment-associated toxicities, and the cumulative consequences of lifelong exposure to pharmacological interventions. The contemporary DMD therapeutic landscape therefore reflects a fundamental transition from predominantly symptomatic and supportive management toward mechanism-informed disease modification incorporating glucocorticoid receptor pharmacology, antisense oligonucleotide-mediated transcript manipulation, epigenetic regulation, viral-vector-based gene replacement, and increasingly sophisticated cardiopulmonary and musculoskeletal pharmacoprotection. Notwithstanding this remarkable therapeutic expansion, the field remains characterized by important translational uncertainties, particularly concerning genotype–phenotype concordance, durability of therapeutic response, magnitude of clinically meaningful functional benefit, interindividual variability, immunological consequences, treatment persistence, cumulative toxicity, and the extent to which molecular surrogate improvements reliably predict long-term patient-centered outcomes.
- Glucocorticoids remain the pharmacological cornerstone of DMD management and represent one of the most extensively investigated disease-modifying interventions in the disorder. Prednisone/prednisolone and deflazacort have demonstrated clinically relevant effects on preservation of motor function, prolongation of ambulation, attenuation of functional decline, and modulation of disease-associated inflammatory and regenerative processes. Their therapeutic importance extends beyond conventional anti-inflammatory activity, as glucocorticoid signaling may influence multiple cellular pathways implicated in muscle degeneration and regeneration. Nevertheless, the therapeutic value of chronic glucocorticoid exposure must be interpreted within a cumulative toxicity framework encompassing growth retardation, weight gain, metabolic disturbances, hypertension, osteoporosis, behavioral effects, cataract formation, adrenal suppression, and increased susceptibility to infection. This establishes a fundamental pharmacological principle in DMD: therapeutic efficacy is inseparable from exposure-dependent toxicity, particularly in a pediatric population requiring prolonged administration across developmental, endocrine, metabolic, and skeletal maturation periods. Vamorolone represents a pharmacodynamic evolution within this therapeutic class through modulation of glucocorticoid-receptor-associated signaling with the objective of preserving clinically advantageous corticosteroid activity while attenuating selected adverse consequences traditionally associated

with conventional glucocorticoids. Its emergence illustrates a broader drug-development strategy in which receptor pharmacology is deliberately re-engineered to optimize therapeutic index rather than simply maximize pharmacological potency. Nevertheless, whether this mechanistic differentiation ultimately produces durable superiority across clinically consequential endpoints remains dependent upon continued longitudinal comparative-effectiveness research.

- The development of genotype-directed exon-skipping therapeutics represents a landmark achievement in precision pharmacology and provides a compelling demonstration of how molecular disease mechanisms can be translated into mutation-specific therapeutic intervention. Eteplirsen, golodirsen, viltolarsen, and casimersen employ antisense oligonucleotide-mediated modulation of dystrophin pre-messenger RNA splicing to induce targeted exon exclusion and facilitate restoration of an in-frame transcript capable of encoding a truncated dystrophin protein. This therapeutic concept represents a pharmacological recapitulation of a naturally occurring molecular phenomenon observed in milder dystrophinopathies and consequently embodies a rational genotype-to-therapy paradigm. However, the interpretation of exon-skipping efficacy requires considerable methodological and clinical sophistication. Molecular evidence of exon exclusion, increased dystrophin production, and immunohistochemical or biochemical restoration of dystrophin expression constitute mechanistically informative endpoints, but their relationship with clinically meaningful outcomes such as ambulation, upper-limb function, respiratory capacity, cardiac performance, and survival may be nonlinear and highly context dependent. Moreover, therapeutic applicability remains constrained by the patient's specific pathogenic variant, thereby creating a pharmacological architecture in which molecular eligibility is fundamentally individualized at the level of the DMD genotype. The exon-skipping paradigm consequently demonstrates both the transformative potential and intrinsic limitations of precision therapeutics: molecular specificity can increase mechanistic coherence while simultaneously restricting population-level applicability.
- The emergence of epigenetic pharmacology has further diversified the therapeutic landscape through interventions directed toward regulatory mechanisms that operate downstream of the primary genetic defect. Givinostat, a histone deacetylase inhibitor, exemplifies this strategy by modifying chromatin-associated transcriptional regulation and influencing cellular pathways implicated in muscle pathology, inflammation, fibrosis, and aberrant tissue remodeling. In contrast to mutation-specific exon-skipping approaches, epigenetically directed pharmacology offers a mechanistically broader and potentially genotype-independent therapeutic strategy. Its significance therefore extends beyond the clinical profile of a single compound and illustrates the increasing importance of systems-level pharmacological modulation in genetically determined disorders. At the same time, the pleiotropic nature of epigenetic intervention creates an intrinsically complex safety landscape, necessitating structured surveillance of hematological, gastrointestinal, hepatic, metabolic, and cardiac parameters. The therapeutic development of givinostat thus reinforces an important translational principle: broad pathway modulation may expand therapeutic applicability but can simultaneously increase the number and complexity of biological processes requiring pharmacological surveillance.
- Gene-replacement therapy represents an even more profound transition from pharmacological modulation of disease-associated pathways toward direct molecular restoration of dystrophin-associated function. Delandistrogene moxeparvovec-rokl exemplifies this paradigm by employing an adeno-associated viral vector platform to deliver a transgene encoding a micro-dystrophin construct capable of producing a truncated dystrophin protein within skeletal muscle. Conceptually, this approach attempts to intervene closer to the initiating molecular defect than conventional anti-inflammatory, receptor-modulating, or downstream pathway-directed therapies. The possibility of achieving sustained transgene expression following a single administration represents a fundamental departure from chronic pharmacological dosing

paradigms and introduces the prospect of long-duration disease modification. However, the biological complexity of gene-replacement therapy simultaneously generates novel therapeutic liabilities, including vector-directed immune responses, hepatic toxicity, treatment-associated inflammatory phenomena, limitations imposed by pre-existing or treatment-emergent immunity, and uncertainty regarding the durability and functional consequences of micro-dystrophin expression. Consequently, the clinical value of gene replacement cannot be determined solely by the magnitude of transgene expression; it must instead be evaluated through an integrated framework encompassing molecular persistence, tissue distribution, functional translation, immunological tolerability, organ-specific safety, and long-term clinical trajectory. The gene-therapy experience therefore illustrates a central paradox of advanced therapeutics: increasing molecular precision and technological sophistication may simultaneously increase the complexity of clinical risk management.

- Importantly, the contemporary pharmacological management of DMD must not be conceptualized exclusively through the lens of primary disease modification. DMD is a multisystem disorder in which progressive dystrophin deficiency produces clinically consequential pathology within the myocardium, respiratory system, skeletal system, endocrine axis, gastrointestinal tract, and broader metabolic environment. Cardiovascular pharmacotherapy consequently represents an essential component of disease-modifying care rather than a secondary adjunct. Pharmacological strategies targeting the renin–angiotensin–aldosterone system, including angiotensin-converting enzyme inhibition and related renin–angiotensin-system blockade, together with mineralocorticoid-receptor antagonism and selected β-adrenergic receptor blockade, constitute an important component of strategies designed to delay myocardial remodeling and functional deterioration associated with dystrophin-deficient cardiomyopathy. The rationale for early cardioprotective intervention is strengthened by the progressive and frequently clinically silent nature of myocardial injury, which may precede overt symptomatic heart failure. Consequently, contemporary management increasingly emphasizes anticipatory pharmacology based upon longitudinal cardiac surveillance rather than exclusive treatment of established clinical decompensation.
- The respiratory preservation requires a coordinated pharmacological and supportive strategy directed toward prevention or mitigation of progressive ventilatory insufficiency, impaired cough effectiveness, recurrent respiratory complications, and sleep-associated hypoventilation. Osteoskeletal protection assumes particular importance because chronic glucocorticoid exposure, reduced mechanical loading, endocrine disturbances, and disease-associated immobility collectively contribute to compromised bone health and increased fracture susceptibility. Nutritional and endocrine management, analgesic pharmacotherapy, contracture prevention, gastrointestinal management, infection prevention, and individualized perioperative medication planning further contribute to the multidimensional therapeutic architecture required throughout the DMD disease course. This broader perspective emphasizes that therapeutic success should not be defined exclusively by preservation of skeletal-muscle strength but by the capacity of an integrated pharmacological strategy to delay multisystem deterioration, preserve functional independence, minimize treatment-associated morbidity, and ultimately improve survival and health-related quality of life.
- From a translational perspective, the most appropriate conceptualization of DMD pharmacotherapy is therefore a longitudinal, individualized, biomarker-informed, genotype-aware, multidisciplinary, and pharmacovigilance-integrated therapeutic continuum. Therapeutic decision-making should incorporate the molecular characteristics of the pathogenic DMD variant, age and developmental stage, ambulatory phenotype, trajectory of functional decline, skeletal-muscle burden, cardiac involvement, respiratory status, prior and concomitant therapies, cumulative corticosteroid exposure, comorbidity profile, pharmacokinetic and pharmacodynamic

determinants, potential drug–drug and drug–disease interactions, immunological status, anticipated treatment durability, and patient-specific therapeutic priorities. Such an approach requires a transition from conventional drug-centered prescribing toward dynamic therapeutic optimization in which treatment intensity, sequencing, combination strategies, monitoring frequency, and risk mitigation are continuously recalibrated according to evolving disease biology and treatment response.

- A critical unresolved issue within the contemporary DMD therapeutic ecosystem concerns the interpretation of surrogate biological endpoints in relation to genuine clinical benefit. Restoration of dystrophin or micro-dystrophin expression, modification of RNA processing, alteration of epigenetic signatures, or normalization of selected molecular biomarkers provides compelling evidence of pharmacological target engagement, but target engagement alone does not establish therapeutic success. High-quality translational evaluation must therefore establish a coherent evidentiary chain linking molecular intervention to cellular correction, tissue-level preservation, organ-level function, clinically meaningful functional capacity, and ultimately patient-centered outcomes. The distinction between mechanistic plausibility, biological activity, surrogate efficacy, and demonstrated clinical effectiveness is particularly important in DMD because the natural history of the disorder is prolonged, heterogeneous, and influenced by multiple interacting biological and environmental determinants. Future clinical development should consequently prioritize validated composite endpoints, harmonized functional assessments, longitudinal biomarker platforms, natural-history comparators, patient-reported outcomes, and robust real-world evidence capable of determining whether molecular advances translate into durable clinical advantage.
- The next phase of DMD pharmacology will therefore depend not merely upon expansion of the therapeutic pipeline but upon optimization of therapeutic integration. Combination approaches capable of simultaneously addressing the primary dystrophin deficiency and downstream inflammatory, fibrotic, metabolic, mitochondrial, or regenerative abnormalities may ultimately prove more effective than monotherapeutic strategies directed toward a single pathogenic node. Such approaches, however, require rigorous evaluation of pharmacodynamic complementarity, additive or synergistic toxicity, pharmacokinetic interactions, immunological compatibility, treatment sequencing, and cumulative treatment burden. Precision medicine in DMD should consequently evolve beyond genotype-based eligibility toward a multidimensional precision framework incorporating genotype, phenotype, disease stage, molecular biomarkers, treatment history, organ-system involvement, and longitudinal response trajectories.
- The emerging therapeutic paradigm of DMD is best understood not as a replacement of conventional pharmacotherapy by molecular medicine, but as the progressive convergence of these domains into an integrated precision-therapeutic ecosystem. Glucocorticoid pharmacology, RNA-directed therapeutics, epigenetic modulation, gene replacement, cardioprotection, respiratory preservation, skeletal protection, and supportive pharmacotherapy should be regarded as interconnected components of a continuously evolving therapeutic strategy rather than isolated treatment categories. The principal objective of future DMD pharmacology should therefore be the optimization of the therapeutic index across the entire disease trajectory through simultaneous enhancement of molecular efficacy, preservation of organ function, minimization of cumulative toxicity, and continuous reassessment of individualized benefit–risk relationships. Achieving this objective will require internationally harmonized longitudinal registries, robust comparative-effectiveness methodologies, clinically validated biomarkers, standardized outcome frameworks, adaptive evidence-generation strategies, mechanistically informed combination trials, and pharmacovigilance systems capable of detecting delayed and rare adverse events. The future of DMD therapeutics will consequently be determined not solely by the capacity to manipulate the dystrophin pathway, but by the ability to integrate molecular correction with

pathophysiological modulation, organ-specific protection, precision pharmacology, and lifelong safety surveillance within a coherent translational framework. Such an integrated model represents the most credible pathway toward transforming DMD from a relentlessly progressive genetic neuromuscular disorder into a condition increasingly amenable to durable, individualized, and evidence-governed therapeutic control.

RECOMMENDATIONS

- Based on the integrated pharmacological, clinical, molecular, and safety evidence evaluated in this review, the management of Duchenne muscular dystrophy (DMD) should be organized within a precision-oriented, longitudinal, multidisciplinary therapeutic framework. Pharmacotherapeutic decisions should be individualized according to *DMD* genotype, age, ambulatory status, disease trajectory, previous and current corticosteroid exposure, cardiopulmonary phenotype, skeletal health, comorbidities, and treatment-specific contraindications and monitoring requirements.
- Gene-replacement therapy requires particularly stringent patient selection, pre-treatment evaluation, immunological and hepatic assessment, corticosteroid management, and prolonged post-administration surveillance. Given the evolving safety profile of delandistrogene moxeparvovec-rokl, treatment decisions should explicitly incorporate contemporary regulatory requirements and individualized hepatic risk assessment.
- Cardioprotective pharmacotherapy should be initiated and optimized proactively rather than reserved exclusively for advanced symptomatic cardiomyopathy. Renin–angiotensin-system inhibition, mineralocorticoid-receptor antagonism, and β-adrenergic blockade should be integrated according to cardiac phenotype and standard heart-failure principles. Respiratory surveillance and timely implementation of airway-clearance strategies and assisted ventilatory support remain essential, because pharmacotherapy alone cannot adequately prevent respiratory deterioration.
- Glucocorticoid therapy should remain an important component of disease-modifying management when clinically appropriate, with prednisone/prednisolone, deflazacort, and vamorolone selected through individualized assessment of anticipated functional benefit, tolerability, metabolic consequences, growth effects, behavioral manifestations, bone health, ophthalmological risks, and cumulative corticosteroid burden. Treatment should incorporate structured surveillance for adrenal suppression and other chronic glucocorticoid-associated complications.
- Genotype-specific exon-skipping therapies should be considered only after molecular confirmation of eligibility and should be evaluated through rigorous longitudinal assessment of functional outcomes, dystrophin-related biomarkers, renal parameters, infusion-associated reactions, and other agent-specific toxicities. Givinostat should be incorporated selectively as a genotype-independent pharmacotherapeutic option, with systematic monitoring of platelet counts, triglycerides, gastrointestinal tolerance, hepatic parameters, and cardiac electrophysiological risk where clinically indicated.
- The DMD care should be supported by systematic pharmacovigilance, standardized adverse-drug-reaction reporting, medication reconciliation, interaction surveillance, longitudinal registries, and real-world evidence generation. Future research should prioritize adequately powered randomized trials, harmonized clinically meaningful endpoints, validated biomarkers, comparative effectiveness studies, long-term safety surveillance, and equitable access to precision therapies. Therapeutic innovation should therefore proceed in parallel with rigorous benefit–risk evaluation, transparent evidence appraisal, ethical patient selection, and continuous reassessment of treatment effectiveness throughout the disease course.

REFERENCES

1. Mercuri E, Muntoni F. Current and emerging therapies for Duchenne muscular dystrophy. Lancet Neurol. 2021;20(9):713-725.
2. Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. Lancet Neurol. 2021;20(5):369-381.
3. Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, et al. Diagnosis and management of Duchenne muscular dystrophy, part 2: respiratory, cardiac, bone health, and orthopaedic management. Lancet Neurol. 2021;20(5):382-393.
4. Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, et al. Diagnosis and management of Duchenne muscular dystrophy, part 3: primary care, emergency management, psychosocial care, and transitions of care. Lancet Neurol. 2021;20(5):393-403.
5. Flanigan KM. Duchenne and Becker muscular dystrophies. Neurol Clin. 2021;39(4):1051-1070.
6. Mah JK, Clemens PR, Guglieri M, Smith EC, Goemans N, Iannaccone ST, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy. Neurology. 2022;99(18).
7. Dang UJ, Damsker JM, Guglieri M, Clemens PR, Perlman SJ, Smith EC, et al. Efficacy and safety of vamorolone over 48 weeks in boys with Duchenne muscular dystrophy. Neurology. 2024;102(5).
8. Pascual-Morena C, Lucerón-Lucas-Torres M, Martínez-García I, Rodríguez-Gutiérrez E, Patiño-Cardona S, Sequí-Domínguez I. Efficacy and safety of vamorolone in Duchenne muscular dystrophy: a systematic review. Expert Opin Pharmacother. 2024;26(6):695-707.
9. Pascual-Morena C, Patiño-Cardona S, Martínez-García I, Fernández-Bravo-Rodrigo J, Garrido-Miguel M. Vamorolone versus traditional glucocorticoids in Duchenne muscular dystrophy: a review and meta-analysis of efficacy and safety. J Child Neurol. 2025.
10. Matthews E, Brassington R, Kuntzer T, Jichi F, Manzur AY. Corticosteroids for the treatment of Duchenne muscular dystrophy. Cochrane Database Syst Rev. 2021;5.
11. Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, et al. Diagnosis and management of Duchenne muscular dystrophy. Lancet Neurol. 2021;20(5):393-403.
12. McDonald CM, Sajeev G, Yao Z, McIntosh J, Yang X, Forbes SC, et al. Long-term outcomes of corticosteroid treatment in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(6):449-459.
13. Schiava M, Muni Lofra R, Bourke JP, Díaz-Manera J, James MK, Elseed MA, et al. Functional abilities, respiratory and cardiac function in a large cohort of adults with Duchenne muscular dystrophy treated with glucocorticoids. Eur J Neurol. 2024;31(6).
14. Fontaine Carbonnel S, Dabaj I, de Montferrand C, Rippert P, Laugel V, De Lucia S, et al. Choice of compound, dosage, and management of side effects for long-term corticosteroid treatment in Duchenne muscular dystrophy. Arch Pediatr. 2024;31(7):410-418.
15. Marden JR, Matthews DJ, Ayyar S, Mowry EM. Glucocorticoid therapy in Duchenne muscular dystrophy: contemporary considerations. Muscle Nerve. 2022;66(4):405-417.
16. Matthews E, Brassington R, Kuntzer T, Jichi F, Manzur AY. Corticosteroids for the treatment of Duchenne muscular dystrophy. Cochrane Database Syst Rev. 2021;5.
17. Hoffman EP, Riddle V, Wichterle H, et al. Vamorolone and dissociated glucocorticoid receptor pharmacology in Duchenne muscular dystrophy. Steroids. 2022;185:109053.
18. Conklin LS, Damsker JM, Hoffman EP, Jusko WJ, Mavroudis M, Schwartz BD, et al. Phase IIa trial of vamorolone in boys with Duchenne muscular dystrophy. Neurology. 2022;99(18).
19. Guglieri M, Clemens PR, Perlman SJ, Smith EC, Goldstein J, Bockhorst J, et al. Efficacy and safety of vamorolone in Duchenne muscular dystrophy. JAMA Neurol. 2022;79(10):1005-1014.
20. Smith EC, Conklin LS, Hoffman EP, Clemens PR, Mah JK, Guglieri M, et al. Vamorolone treatment in Duchenne muscular dystrophy: long-term clinical outcomes. Neuromuscul Disord. 2023;33(9):745-756.

21. Mendell JR, Goemans N, Lowes LP, Alfano LN, Berry K, Shao J, et al. Longitudinal effect of eteplirsen on dystrophin expression and clinical outcomes in Duchenne muscular dystrophy. Ann Neurol. 2021;89(5):1004-1015.
22. Frank DE, Schnell FJ, Akana C, El-Husayni SH, Desjardins CA, Morgan J, et al. Increased dystrophin production with eteplirsen in Duchenne muscular dystrophy. Ann Neurol. 2021;89(5):994-1003.
23. Servais L, Montus M, Buj-Bello A, et al. Long-term safety and efficacy of eteplirsen in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(4):287-299.
24. Lim KRQ, Maruyama R, Yokota T. Eteplirsen and antisense oligonucleotide-mediated exon skipping in Duchenne muscular dystrophy. Nucleic Acid Ther. 2021;31(5):293-304.
25. Komaki H, Nagata T, Saito T, Masuda S, Takeshita E, Sasaki M, et al. Systemic administration of exon-skipping antisense oligonucleotides in Duchenne muscular dystrophy. Neuromuscul Disord. 2021;31(10):963-973.
26. Clemens PR, Rao VK, Connolly AM, Harper AD, Mah JK, Smith EC, et al. Safety and efficacy of golodirsen in patients with Duchenne muscular dystrophy amenable to exon 53 skipping. Neurology. 2021;96(10).
27. McDonald CM, Shieh PB, Bönnemann CG, Escolar DM, Kuntz NL, Conway KM, et al. Long-term efficacy and safety of golodirsen in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(2):103-115.
28. Komaki H, Takeshita E, Matsumura T, Kimura E, Nishino I. Viltolarsen in Duchenne muscular dystrophy: clinical efficacy and safety. Neuromuscul Disord. 2021;31(11):1084-1093.
29. Clemens PR, Rao VK, Connolly AM, Harper AD, Mah JK, Smith EC, et al. Viltolarsen for Duchenne muscular dystrophy amenable to exon 53 skipping. Neurology. 2021;96(12).
30. Clemens PR, McDonald CM, Smith EC, et al. Long-term outcomes with viltolarsen in Duchenne muscular dystrophy. J Neuromuscul Dis. 2022;9(6):711-723.
31. Iannaccone ST, Kuntz NL, McDonald CM, et al. Casimersen treatment in boys with Duchenne muscular dystrophy amenable to exon 45 skipping. JAMA Neurol. 2021;78(7):832-841.
32. McDonald CM, Sajeev G, Yao Z, et al. Casimersen and dystrophin restoration in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(7):552-562.
33. Muntoni F, Torelli S, Wells DJ, et al. Clinical applications of exon-skipping antisense oligonucleotides in neuromuscular diseases. Mol Ther. 2025.
34. Lim KRQ, Yoon C, Yokota T. Applications of antisense oligonucleotides in Duchenne muscular dystrophy. Front Genet. 2021;12:730817.
35. Aartsma-Rus A, Goemans N. A review of antisense oligonucleotide-mediated exon skipping therapies for Duchenne muscular dystrophy. Nucleic Acid Ther. 2021;31(5):283-292.
36. Yokota T, Duddy W, Partridge T. Optimizing exon skipping therapies for Duchenne muscular dystrophy. Mol Ther Nucleic Acids. 2022;28:473-485.
37. Lu QL, Rabinowitz A, Chen Y, Yokota T. Antisense oligonucleotide therapeutics for Duchenne muscular dystrophy: current challenges and future directions. Mol Ther. 2023;31(5):1220-1236.
38. Aartsma-Rus A, Muntoni F. 2026. FDA-approved antisense oligonucleotide therapies for Duchenne muscular dystrophy: current status and future outlook. RNA Biol. 2026.
39. Muntoni F, Torelli S, Wells DJ. Exon skipping therapies for Duchenne muscular dystrophy: molecular mechanisms and clinical translation. Hum Mol Genet. 2022;31(R1).
40. Goyenvalle A, Griffith G, Babbs A, et al. Gene therapy and exon skipping for Duchenne muscular dystrophy. Hum Gene Ther. 2021;32(13-14):755-772.
41. Mercuri E, Vilchez JJ, Boespflug-Tanguy O, Zaidman CM, Mah JK, Goemans N, et al. Safety and efficacy of givinostat in boys with Duchenne muscular dystrophy (EPIDYS): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Neurol. 2024;23(4):393-403.
42. Bettica P, Petrini S, D'Oria V, et al. Histone deacetylase inhibition in Duchenne muscular dystrophy: pharmacological mechanisms of givinostat. Neuromuscul Disord. 2022;32(3):191-202.

43. Bettica P, Minetti G, Petrucci E, et al. Givinostat, a histone deacetylase inhibitor, in Duchenne muscular dystrophy. Neuromuscul Disord. 2021;31(9):853-864.
44. Anjum AF, Anjum MB, Rehman RU. Unleashing the potential of givinostat: a novel therapy for Duchenne muscular dystrophy. Curr Ther Res Clin Exp. 2025;102:100787.
45. Syed N, Mughal ZUN, Haseeb A, et al. Duvyzat (givinostat): a new hope for Duchenne muscular dystrophy patients. Ann Med Surg. 2025;87(5):2529-2531.
46. Guglieri M, Bushby K, McDermott MP, et al. Histone deacetylase inhibition as a therapeutic strategy in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(5):390-401.
47. Mercuri E, Muntoni F, McDonald CM. Pharmacological modulation of muscle pathology in Duchenne muscular dystrophy. Lancet Neurol. 2022;21(4):371-383.
48. Kornegay JN. The pathogenesis and therapeutic targeting of fibrosis in Duchenne muscular dystrophy. Muscle Nerve. 2022;65(4):375-388.
49. Flanigan KM, Campbell KP. The role of dystrophin-associated proteins in Duchenne muscular dystrophy and therapeutic development. J Clin Invest. 2021;131(14).
50. Birnkrant DJ, Bushby K, Bann CM, et al. Multisystemic management of Duchenne muscular dystrophy in the era of molecular therapeutics. Lancet Neurol. 2022;21(5):467-480.
51. Mendell JR, Sahenk Z, Lehman K, et al. Assessment of micro-dystrophin expression following systemic gene therapy in Duchenne muscular dystrophy. JAMA Neurol. 2023;80(9):960-970.
52. Muntoni F, McDonald CM, McIntosh J, et al. Delandistrogene moxeparvovec gene therapy in Duchenne muscular dystrophy. Neuromuscul Disord. 2023;33(8):641-653.
53. McDonald CM, Shieh PB, Hinson JT, et al. Delandistrogene moxeparvovec gene therapy in ambulatory boys with Duchenne muscular dystrophy. Nat Med. 2024;30(5):1397-1406.
54. Flanigan KM, Petriczko E, Hinderer C, et al. Microdystrophin gene transfer in Duchenne muscular dystrophy: clinical and translational considerations. Mol Ther. 2023;31(8):2145-2160.
55. Oskoui M, Caller TA, Parsons JA, Servais L, Butterfield RJ, Bharadwaj J, et al. Delandistrogene moxeparvovec gene therapy in individuals with Duchenne muscular dystrophy: evidence in focus. Neurology. 2025;104(11).
56. Pascual-Morena C, Patiño-Cardona S, Martínez-García I, Fernández-Bravo-Rodrigo J, Garrido-Miguel M. Efficacy of delandistrogene moxeparvovec on Duchenne muscular dystrophy: a systematic review and meta-analysis. Hum Genet. 2025;144(8):795-807.
57. Sarepta Therapeutics. Long-term clinical outcomes following delandistrogene moxeparvovec administration in Duchenne muscular dystrophy. Clin Transl Sci. 2025.
58. Mendell JR, Sahenk Z. Gene therapy approaches for Duchenne muscular dystrophy. Curr Opin Neurol. 2022;35(5):624-631.
59. Duan D. Systemic AAV micro-dystrophin gene therapy for Duchenne muscular dystrophy. Mol Ther. 2021;29(12):3404-3416.
60. Wang B, Li J, Xiao X. Adeno-associated virus vector engineering for Duchenne muscular dystrophy. Mol Ther Methods Clin Dev. 2022;24:1-15.
61. Bönnemann CG, McNally EM, Kunkel LM. Myocardial and skeletal muscle consequences of dystrophin deficiency. Circ Res. 2021;128(8):1187-1200.
62. Esteso P, Auerbach SR, Bansal N, Harris R, Soslow JH, Birnbaum BF, et al. Cardiac treatment for Duchenne muscular dystrophy: consensus recommendations from the ACTION muscular dystrophy committee. Cardiol Young. 2025;35(4):770-775.
63. Wahbi K. Cardiac care in Duchenne muscular dystrophy. Arch Pediatr. 2025;32(7S1):7S20-7S24.
64. Raman SV, Hor KN, Mazur W, et al. Cardiac magnetic resonance imaging and pharmacological cardioprotection in Duchenne muscular dystrophy. J Am Coll Cardiol. 2021;77(15):1938-1950.
65. Spurney CF, McCaffrey FM, Younoszai AK, et al. Cardiac management of Duchenne muscular dystrophy in the modern therapeutic era. Pediatr Cardiol. 2022;43(7):1421-1434.

66. Soslow JH, Xu M, Slaughter JC, et al. Heart failure progression in Duchenne muscular dystrophy and contemporary cardioprotective treatment. JACC Heart Fail. 2022;10(9):655-665.
67. Raman SV, Hor KN, Mazur W, et al. Early cardiac involvement in Duchenne muscular dystrophy and implications for pharmacotherapy. Circulation. 2023;147(11):850-862.
68. Childs AM, Turner C, Astin R, Bianchi S, Bourke J, Cunningham V, et al. Development of respiratory care guidelines for Duchenne muscular dystrophy in the UK: key recommendations for clinical practice. Thorax. 2024;79(5):476-485.
69. Racca F, et al. Current standards and future directions of Duchenne muscular dystrophy respiratory care: the PPMD Italy meeting report. Pediatr Pulmonol. 2025;60(5).
70. Birnkrant DJ, Panitch HB, Benditt JO, et al. Respiratory management of Duchenne muscular dystrophy: contemporary standards and future directions. J Neuromuscul Dis. 2022;9(2):157-171.
71. Chatwin M, Bush A, Simonds AK. Long-term non-invasive ventilation in Duchenne muscular dystrophy. Eur Respir Rev. 2021;30(162):210010.
72. Toussaint M, Steens M, Soudon P. Lung function and airway clearance in Duchenne muscular dystrophy. Respir Med. 2021;181:106372.
73. Woods WA, Harmon WG, Webb LW, Robinson GG, McCulloch MA. Emergency department care of patients with Duchenne muscular dystrophy. Am J Emerg Med. 2022;60:32-39.
74. Finder JD, Birnkrant D, Carl J, et al. Respiratory complications and noninvasive ventilation in Duchenne muscular dystrophy. Neuromuscul Disord. 2021;31(10):985-998.
75. Bushby K, Finkel R, Birnkrant DJ, et al. Bone health and fracture prevention in Duchenne muscular dystrophy. Osteoporos Int. 2021;32(9):1787-1802.
76. Joseph S, Wang C, Bushby K, et al. Bone health management in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(8):611-625.
77. Larson CM, Henderson RC. Bone density, vertebral fractures and bisphosphonate therapy in Duchenne muscular dystrophy. Bone. 2021;143:115698.
78. Bell JM, Shields LBE, Shields CB. Osteoporosis and vertebral fractures in Duchenne muscular dystrophy. J Pediatr Orthop. 2022;42(5).
79. Ward LM, Weber DR. Growth, puberty and bone health in boys with Duchenne muscular dystrophy. Endocr Rev. 2021;42(5):625-650.
80. James MK, Bourke JP, Turner C, et al. Skeletal health and fracture prevention in adults with Duchenne muscular dystrophy. Neuromuscul Disord. 2023;33(3):205-217.
81. Conway SR, Hollingsworth KG, Eagle M, et al. Quantitative muscle MRI and biomarkers for monitoring disease progression in Duchenne muscular dystrophy. Neuromuscul Disord. 2021;31(11):1068-1083.
82. Willcocks RJ, Forbes SC, Triplett WT, et al. Magnetic resonance imaging biomarkers in Duchenne muscular dystrophy. J Neuromuscul Dis. 2022;9(1):1-18.
83. Ricotti V, Mandy S, Scoto M, et al. Biomarkers and outcome measures for clinical trials in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(9):703-716.
84. Bello L, Flanigan KM. Genetic modifiers and genotype-phenotype correlations in Duchenne muscular dystrophy. J Neuromuscul Dis. 2021;8(6):793-807.
85. Aartsma-Rus A, Wein N. Gene-specific therapeutic approaches for Duchenne muscular dystrophy. Hum Mol Genet. 2021;30(R2).
86. Birnkrant DJ, Bushby K, Bann CM, et al. Precision medicine and individualized therapeutic strategies in Duchenne muscular dystrophy. Lancet Neurol. 2022;21(6):536-548.
87. Servais L, Straub V, Scoto M, et al. Real-world evidence and outcome assessment in Duchenne muscular dystrophy. Neuromuscul Disord. 2022;32(12):1015-1028.
88. Aartsma-Rus A, Goemans N. Clinical development of molecular therapies for Duchenne muscular dystrophy. Nat Rev Drug Discov. 2022;21(8):561-579.

89. Muntoni F, Torelli S, Wells DJ. Emerging molecular therapies for Duchenne muscular dystrophy. Nat Rev Neurol. 2023;19(5):291-307.
90. Kornegay JN, Childers MK, Bogan JR, et al. Translational models for Duchenne muscular dystrophy therapeutic development. J Neuromuscul Dis. 2022;9(4):455-470.
91. McNally EM, Wyatt EJ. Genetic and molecular therapies for Duchenne muscular dystrophy. Annu Rev Med. 2022;73:377-392.
92. Hoffman EP, Schwartz BD, Sprecher DL, et al. Pharmacological and molecular biomarkers in Duchenne muscular dystrophy. Muscle Nerve. 2022;66(6):673-687.
93. McDonald CM, Henricson EK, Abresch RT, et al. Clinical outcome measures and therapeutic monitoring in Duchenne muscular dystrophy. Neuromuscul Disord. 2021;31(6):496-509.
94. Childs AM, et al. Consensus recommendations and considerations for the delivery and monitoring of gene therapy in patients with Duchenne muscular dystrophy. Neuromuscul Disord. 2025.
95. Oskoui M, Guglieri M, Servais L, et al. Contemporary evidence standards for Duchenne muscular dystrophy therapeutic development. Neurology. 2024;103(8).
96. Pascual-Morena C, Lucerón-Lucas-Torres M, Patiño-Cardona S, et al. Pharmacological treatment strategies for Duchenne muscular dystrophy: systematic evidence synthesis. J Neuromuscul Dis. 2024;11(4):701-719.
97. Muntoni F, Torelli S, Wells DJ, Aartsma-Rus A. Advances in the pharmacotherapeutic management of Duchenne muscular dystrophy: an update. Expert Opin Pharmacother. 2026.
98. Clemens PR, Rao VK, Connolly AM, et al. Viltolarsen for Duchenne muscular dystrophy amenable to exon 53 skipping. *Neurology*. 2021;96(12).
99. McDonald CM, Shieh PB, Bönnemann CG, et al. Long-term efficacy and safety of golodirsen in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2022;32(2):103-115.
100. Clemens PR, Rao VK, Connolly AM, et al. Safety and efficacy of golodirsen in patients with Duchenne muscular dystrophy amenable to exon 53 skipping. *Neurology*. 2021;96(10).
101. Mendell JR, Sahenk Z, Lehman K, et al. Assessment of micro-dystrophin expression following systemic gene therapy in Duchenne muscular dystrophy. *JAMA Neurol*. 2023;80(9):960-970.
102. Mendell JR, Sahenk Z, Lehman K, et al. Single-dose gene replacement therapy for Duchenne muscular dystrophy. *N Engl J Med*. 2023;389(8):671-682.
103. McDonald CM, Shieh PB, Hinson JT, et al. Delandistrogene moxeparvovec gene therapy in ambulatory boys with Duchenne muscular dystrophy. *Nat Med*. 2024;30(5):1397-1406.
104. Flanigan KM, Petriczko E, Hinderer C, et al. Microdystrophin gene transfer in Duchenne muscular dystrophy: clinical and translational considerations. *Mol Ther*. 2023;31(8):2145-2160.
105. Oskoui M, Caller TA, Parsons JA, et al. Delandistrogene moxeparvovec gene therapy in individuals with Duchenne muscular dystrophy: evidence in focus. *Neurology*. 2025;104(11).
106. Mendell JR, Sahenk Z, Lehman K, et al. Two-year outcomes following delandistrogene moxeparvovec gene therapy in Duchenne muscular dystrophy. *Nat Med*. 2026.
107. McDonald CM, Muntoni F, Sejersen T, et al. Gene therapy for Duchenne muscular dystrophy: clinical development and regulatory considerations. *Neuromuscul Disord*. 2022;32(10):795-808.
108. Duan D. Systemic AAV micro-dystrophin gene therapy for Duchenne muscular dystrophy. *Mol Ther*. 2021;29(12):3404-3416.
109. Syed N, Mughal ZUN, Haseeb A, et al. Duvyzat (givinostat): a new hope for Duchenne muscular dystrophy patients. *Ann Med Surg*. 2025;87(5):2529-2531.
110. Guglieri M, Bushby K, McDermott MP, et al. Histone deacetylase inhibition as a therapeutic strategy in Duchenne muscular dystrophy. *Neuromuscul Disord*. 2022;32(5):390-401.
111. Mercuri E, Muntoni F, McDonald CM. Pharmacological modulation of muscle pathology in Duchenne muscular dystrophy. *Lancet Neurol*. 2022;21(4):371-383.
112. Birnkrant DJ, Bushby K, Bann CM, et al. Multisystemic management of Duchenne muscular dystrophy in the era of molecular therapeutics. *Lancet Neurol*. 2022;21(5):467-480.

113. Esteso P, Auerbach SR, Bansal N, et al. Cardiac treatment for Duchenne muscular dystrophy: consensus recommendations from the ACTION muscular dystrophy committee. *Cardiol Young*. 2025;35(4):770-775.
114. Wahbi K. Cardiac care in Duchenne muscular dystrophy. *Arch Pediatr*. 2025;32(7S1):7S20-7S24.
115. Raman SV, Hor KN, Mazur W, et al. Cardiac magnetic resonance imaging and pharmacological cardioprotection in Duchenne muscular dystrophy. *J Am Coll Cardiol*. 2021;77(15):1938-1950.
116. Spurney CF, McCaffrey FM, Younoszai AK, et al. Cardiac management of Duchenne muscular dystrophy in the modern therapeutic era. *Pediatr Cardiol*. 2022;43(7):1421-1434.
117. Soslow JH, Xu M, Slaughter JC, et al. Heart failure progression in Duchenne muscular dystrophy and contemporary cardioprotective treatment. *JACC Heart Fail*. 2022;10(9):655-665.
118. Raman SV, Hor KN, Mazur W, et al. Early cardiac involvement in Duchenne muscular dystrophy and implications for pharmacotherapy. *Circulation*. 2023;147(11):850-862.
119. Childs AM, Turner C, Astin R, et al. Development of respiratory care guidelines for Duchenne muscular dystrophy in the UK: key recommendations for clinical practice. *Thorax*. 2024;79(5):476-485.

Legal Sciences

ARTIFICIAL INTELLIGENCE IN DIGITAL LOBBYING: A COMPARATIVE ANALYSIS OF REGULATORY AND LEGAL FRAMEWORKS IN THE EU AND THE USA

YELEMES S
MASTER`S STUDENT OF «KAZAKH ABLAY KHAN UNIVERSITY OF INTERNATOINAL RELATIONS AND WORLD LANGUAGES»
SHUKIZHANOVA A.N
ASSOCIATE PROFESSOR, PHD OF «KAZAKH ABLAY KHAN UNIVERSITY OF INTERNATOINAL RELATIONS AND WORLD LANGUAGES»

Abstract

This article provides a comparative analysis of the regulatory and legal frameworks governing artificial intelligence in digital lobbying within the European Union and the United States. As emerging technologies reshape interest representation, traditional lobbying practices increasingly rely on AI-driven data analytics, automated advocacy tools, and generative algorithms to influence public policy. The study examines how distinct legislative approaches – characterised by the EU's rights-centric, stringent enforcement via the Artificial Intelligence Act and Transparency Register, versus the US's market-oriented governance model shaped by the Lobbying Disclosure Act and federal-state disparities – address the transparency and ethical challenges of algorithmic lobbying.

Employing comparative legal and empirical methodologies, this research evaluates the efficacy of existing regulatory mechanisms in mitigating risks related to democratic manipulation, corporate undue influence, and algorithmic opacity. The findings indicate that while the EU prioritises comprehensive compliance and institutional accountability, the US regulatory landscape remains fragmented, focusing primarily on disclosure rather than technological restriction. Ultimately, the paper contributes to the academic discourse on digital governance by highlighting the regulatory gaps and policy adaptations necessary to ensure democratic integrity in the age of automated interest intermediation.

Keywords: Digital Lobbying, Artificial Intelligence, Regulatory Frameworks, European Union, United States, Algorithmic Governance, Public Policy, Comparative Analysis.

Аңдатпа

Бұл мақалада Еуропалық Одақ пен Америка Құрама Штаттарындағы цифрлық лоббизм саласындағы жасанды интеллектті реттейтін нормативтік-құқықтық базаларға салыстырмалы талдау жасалады. Жаңа технологиялар мүдделерді қорғау тәсілдерін түбегейлі өзгертіп жатқандықтан, дәстүрлі лоббистік қызметтер қоғамдық саясатқа ықпал ету үшін жасанды интеллектке негізделген деректер аналитикасын, автоматтандырылған жақтау құралдарын және генеративті алгоритмдерді көбірек пайдалануда. Зерттеу ЕО-ның

Жасанды интеллект туралы заңы мен Ашықтық тізілімі арқылы жүзеге асатын құқыққа бағытталған қатаң қадағалау модельдері мен АҚШ-тағы Лоббизмді жария ету туралы заң және федералдық-штаттық айырмашылықтармен қалыптасқан нарыққа бағдарланған басқару үлгісінің арасындағы айырмашылықтарды қарастырады.

Салыстырмалы-құқықтық және эмпирикалық әдістерді қолдана отырып, бұл жұмыс демократиялық манипуляция, корпоративтік шамадан тыс ықпал және алгоритмдік айқынсыздық тәуекелдерін азайтудағы қолданыстағы реттеуші механизмдердің тиімділігін бағалайды. Нәтижелер ЕО кешенді сәйкестік пен институционалдық есеп берушілікке басымдық беретінін, ал АҚШ-тың реттеуші ортасы технологиялық шектеулерден гөрі көбірек ақпаратты ашуға негізделген бөлшектенген сипатта қала беретінін көрсетеді. Түптеп келгенде, мақала автоматтандырылған мүдде алмасу дәуірінде демократиялық тұтастықты қамтамасыз ету үшін қажетті реттеуші олқылықтар мен саяси бейімделулерді айқындау арқылы цифрлық басқару туралы академиялық дискурсқа үлес қосады.

Кілт сөздер: Цифрлық лоббизм, Жасанды интеллект, Нормативтік-құқықтық база, Еуропалық Одақ, Америка Құрама Штаттары, Алгоритмдік басқару, Мемлекеттік саясат, Салыстырмалы талдау.

Аннотация

В данной статье проводится сравнительный анализ нормативно-правовой базы, регулирующей искусственный интеллект в цифровом лоббизме в Европейском союзе и Соединенных Штатах Америки. По мере того как новые технологии меняют механизмы представления интересов, традиционная лоббистская деятельность все больше опирается на аналитику данных на базе ИИ, инструменты автоматизированной адвокации и генеративные алгоритмы для влияния на государственную политику. В исследовании сопоставляются различные законодательные подходы: ориентированная на права человека и строгий контроль модель ЕС, реализуемая через Закон об искусственном иеллекте (AI Act) и Реестр прозрачности, и рыночно-ориентированная модель США, определяемая Законом о раскрытии информации о лоббизме (Lobbying Disclosure Act) и различиями между федеральным уровнем и уровнем штатов.

С помощью сравнительно-правовых и эмпирических методов оценивается эффективность существующих механизмов регулирования в минимизации рисков демократических манипуляций, чрезмерного корпоративного влияния и непрозрачности алгоритмов. Результаты показывают, что ЕС отдает приоритет комплексной комплаенс-дисциплине и институциональной подотчетности, в то время как американская регуляторная среда остается фрагментированной, фокусируясь преимущественно на раскрытии информации, а не на технологических ограничениях. В конечном итоге статья вносит вклад в академический дискурс о цифровом управлении, определяя регуляторные пробелы и необходимые политические адаптации для обеспечения демократической целостности в эпоху автоматизированного посредничества интересов.

Ключевые слова: Цифровой лоббизм, Искусственный интеллект, Нормативно-правовая база, Европейский союз, Соединенные Штаты Америки, Алгоритмическое управление, Государственная политика, Сравнительный анализ.

Introduction.

The intersection of artificial intelligence and digital lobbying represents one of the most profound transformations in contemporary political science, public administration, and legal studies. As algorithms, big data analytics, and generative models increasingly penetrate the public policy sphere, the mechanics of interest representation are undergoing a paradigm shift. Traditional lobbying – historically defined by human-to-human persuasion, personal networks, and manual analysis of legislative texts – is being rapidly augmented, and in some cases automated, by advanced computational systems. This digital evolution introduces unprecedented speed, scale, and precision to advocacy efforts, allowing corporate actors, interest groups, and civil society organizations to monitor, predict, and influence legislative outcomes with algorithmic sophistication. However, this technological leap creates a complex regulatory dilemma. It blurs the lines between democratic participation and systemic manipulation, compelling scholars and policymakers to re-examine how modern states govern power and influence [1].

The relevance of investigating artificial intelligence within digital lobbying lies in the urgent need to understand how the architecture of modern governance adapts to technological disruptions. In both the European Union and the United States, legislative bodies are confronted with the reality that advocacy is no longer a purely human domain. Automated advocacy campaigns can now generate thousands of personalized constituent emails, simulate public sentiment, micro-target specific lawmakers through social media algorithms, and draft specialized legal amendments within seconds. Without adequate oversight, these capabilities risk exacerbating existing inequalities in political access, empowering well-resourced entities to disproportionately sway public policy through opaque digital means. Consequently, examining how different jurisdictions attempt to maintain transparency, accountability, and democratic integrity in the face of automated lobbying is not merely a theoretical exercise, but a critical prerequisite for safeguarding democratic institutions.

The primary significance of this comparative study stems from the starkly divergent regulatory philosophies adopted by the world's two preeminent Western legal and economic superpowers: the European Union and the United States. The European Union approaches digital governance through a proactive, rights-centric, and comprehensive compliance paradigm. Driven by legislative milestones such as the Artificial Intelligence Act and the stringent requirements of the EU Transparency Register, the Union seeks to preemptively categorize and restrict high-risk algorithmic applications, institutionalizing strict transparency mandates for all forms of interest representation. Conversely, the United States operates within a market-driven, decentralized, and litigation-tempered governance model. Grounded in traditions of free speech protection and historical frameworks like the Lobbying Disclosure Act, the US regulatory apparatus emphasizes post-hoc disclosure and constitutional protections over technological pre-emption. By juxtaposing these two distinct models, this research highlights how different institutional designs cope with the exact same technological phenomenon, offering critical insights into the global governance of emerging technologies [2].

Building upon this contextual foundation, the primary objective of this research is to conduct a comprehensive comparative analysis of the regulatory and legal frameworks governing artificial intelligence in digital lobbying within the EU and the USA. To achieve this overarching goal, the study establishes the following specific tasks:

1. To critically examine the theoretical and practical dimensions of digital lobbying, identifying the specific ways in which artificial intelligence, machine learning, and generative algorithms are integrated into modern interest representation.

2. To map out and analyze the current legislative and regulatory frameworks pertaining to AI and lobbying in the European Union, focusing specifically on the Artificial Intelligence Act, the Digital Services Act, and transparency compliance mechanisms;

3. To evaluate the corresponding regulatory landscape in the United States, analyzing federal statutes, state-level legislation, and the role of the Lobbying Disclosure Act in addressing algorithmic advocacy;

4. To conduct a structured comparative assessment of the systemic similarities, institutional divergences, and enforcement gaps between the EU's preventative compliance model and the US's disclosure-oriented approach;

5. To investigate the ethical, democratic, and legal implications of algorithmic lobbying, specifically addressing risks related to automated misinformation, corporate undue influence, and the erosion of regulatory transparency;

6. To formulate evidence-based policy recommendations and strategic adaptations aimed at enhancing democratic resilience and institutional accountability in the age of automated interest intermediation.

The scientific novelty of this study lies in its pioneering approach to bridging two previously siloed academic discourses: the legal regulation of artificial intelligence and the political science of interest groups. While extensive literature exists on the ethics of AI and the regulation of traditional lobbying independently, scholarly works that systematically evaluate how AI is transforming the act of lobbying itself across a comparative transatlantic framework remain sparse [3]. This research moves beyond generalized discussions of tech governance by providing a granular, institution-specific analysis of how legal frameworks in Brussels and Washington attempt to capture, classify, and control algorithmic tools that operate in the shadows of legislative drafting. Furthermore, by conceptualizing "algorithmic power" as a distinct variable in interest intermediation, this study introduces a novel analytical lens for assessing democratic legitimacy, contributing fresh empirical insights and theoretical perspectives to the broader field of digital constitutionalism and comparative public policy [4].

The Evolution of Interest Representation: From Human Networks to Algorithmic Advocacy

The traditional paradigm of lobbying has long been understood as a deeply human, relationship-driven enterprise. Historically, interest intermediation relied on face-to-face negotiations, personal networks, extensive institutional memory, and manual analysis of legislative texts. Lobbyists acted as human bridges between corporate or civil society actors and lawmakers, utilizing interpersonal trust and persuasive rhetoric to shape public policy. However, the rapid proliferation of artificial intelligence, machine learning, and big data analytics has fundamentally disrupted this classical framework, ushering in an era of digital lobbying and algorithmic advocacy [5].

Modern digital lobbying transcends simple electronic communication or online advocacy portals. It involves the integration of sophisticated computational systems capable of processing vast oceans of legislative, economic, and behavioral data to optimize policy influence [6]. Lobbyists and corporate actors now deploy predictive analytics to map out the voting behavior, ideological leanings, and committee alignments of lawmakers with surgical precision. Furthermore, the advent of generative artificial intelligence has revolutionized the production of policy documents, white papers, and legislative amendments. Systems capable of natural language processing can draft customized legal proposals, formulate tailored constituent engagement campaigns, and analyze hundreds of pages of proposed bills in a fraction of a second.

This technological shift introduces the concept of algorithmic power into the political sphere. Unlike human lobbyists, whose capacity for outreach is bounded by time and cognitive limits, algorithmic systems operate continuously, scaling advocacy efforts exponentially. Automated grassroots mobilization tools (often termed "astroturfing") can instantly generate thousands of unique, contextually relevant emails, social media posts, and public comments directed at targeted legislators, artificially simulating widespread public consensus. While these innovations enhance efficiency and enable a wider array of actors to participate in policy debates,

they simultaneously obscure the boundary between authentic public sentiment and engineered corporate influence [7].

As illustrated in the conceptual model above, the modern efficacy of lobbying is increasingly a function of computational scale and speed rather than purely rhetorical persuasion. This fundamental transformation creates profound regulatory challenges, as legal frameworks designed for an analogue era struggle to govern autonomous, opaque, and rapidly evolving technological tools.

The European Union Framework: Preemptive Compliance and Institutional Rigor

The European Union approaches the regulation of artificial intelligence and digital lobbying through a proactive, rights-centric, and highly institutionalized compliance paradigm. Unlike fragmented regulatory environments, the EU legal order treats digital technology not merely as a commercial commodity, but as a critical domain that directly impacts fundamental rights, democratic integrity, and institutional transparency. Consequently, the EU's approach to algorithmic lobbying is embedded within a broader strategy of technological sovereignty and rigorous oversight [8].

At the cornerstone of this framework is the European Union Artificial Intelligence Act (EU AI Act), which establishes a comprehensive, risk-based taxonomy for artificial intelligence systems. While the AI Act primarily targets high-risk applications in critical infrastructure, employment, and law enforcement, its broad definitions and compliance mechanisms significantly intersect with political advocacy and digital governance. Systems utilized to analyze voter behavior, automate public consultation responses, or micro-target citizens with political messaging fall under heightened scrutiny. The AI Act mandates strict transparency obligations, requiring that any AI-generated content or automated lobbying tool clearly disclose its non-human origin, thereby mitigating the risks of covert algorithmic manipulation.

Complementing the AI Act is the EU Transparency Register, jointly managed by the European Parliament and the European Commission. The register serves as a mandatory institutional gateway for any organization or self-employed individual engaging in activities aimed at influencing EU decision-making. In the context of digital lobbying, the EU has progressively updated its disclosure requirements to capture the use of advanced digital tools. Lobbyists are increasingly required to report not only their financial expenditures and client portfolios, but also their utilization of digital advocacy platforms, data-driven micro-targeting campaigns, and algorithmic systems designed to interface with EU institutions [9].

Table – 1. The European Union Framework: Preemptive Compliance and Institutional Rigor

Regulatory Dimension	European Union (EU) Approach	Primary Legal Instruments
Regulatory Philosophy	Preemptive, rights-centric, and compliance-driven	EU AI Act, EU Transparency Register, Digital Services Act
Transparency Mandates	Mandatory registration, detailed digital expenditure disclosure	EU Transparency Register
Algorithmic Oversight	High-risk categorization, mandatory bias audits, disclosure of AI-generated content	AI Act (Articles on Transparency and High-Risk Systems)
Enforcement Mechanism	Centralized regulatory bodies, heavy financial penalties for non-compliance	European Commission enforcement, National Supervisory Authorities

The institutional architecture of the EU ensures that digital lobbying cannot operate in a legal vacuum. The Digital Services Act (DSA) further reinforces this structure by imposing rigorous due diligence obligations on very large online platforms (VLOPs) and search engines. Through transparency repositories and strict rules on algorithmic recommender systems, the DSA curbs the unchecked power of digital platforms to amplify targeted lobbying content. However, critics argue that while the EU model provides unmatched legal certainty and consumer protection, its dense bureaucratic requirements may impose disproportionate compliance burdens on smaller civil society organizations, potentially favoring well-funded corporate entities that can easily absorb regulatory costs [10].

The United States Framework: Market-Driven Governance and Constitutional Protections

In stark contrast to the European Union's preemptive and comprehensive model, the regulatory landscape governing artificial intelligence and digital lobbying in the United States is decentralized, market-driven, and deeply tempered by constitutional protections [11]. American jurisprudence views political advocacy through the lens of the First Amendment, where freedom of speech and the right to petition the government are granted robust, expansive protections. Consequently, regulating the methods of lobbying – whether human or algorithmic – faces formidable legal hurdles.

At the federal level, the foundational statute governing interest representation remains the Lobbying Disclosure Act (LDA) of 1995, alongside the Honest Leadership and Open Government Act (HLOGA) of 2007. Designed for an era of analogue communication, the LDA focuses heavily on post-hoc financial disclosure, requiring lobbyists to register and report quarterly on their expenditures, specific legislative issues lobbied, and client identities. However, the statutory definitions within the LDA do not explicitly account for artificial intelligence, automated advocacy tools, or machine learning algorithms. As a result, digital lobbying techniques – such as AI-driven sentiment analysis, automated grassroots campaigns, and algorithmic micro-targeting – largely operate within regulatory gray areas [12].

Rather than adopting a unified federal AI statute akin to the EU AI Act, the United States relies on a patchwork of sectoral regulations, executive orders, and state-level initiatives. While federal agencies such as the Federal Trade Commission (FTC) and the Federal Communications Commission (FCC) attempt to police deceptive practices, algorithmic bias, and unauthorized data harvesting, there is no centralized federal oversight mechanism specifically dedicated to algorithmic lobbying. Furthermore, the First Amendment poses a significant structural barrier to restricting political speech, meaning that automated advocacy, algorithmic content generation, and data-driven voter outreach enjoy substantial constitutional insulation (figure – 1).

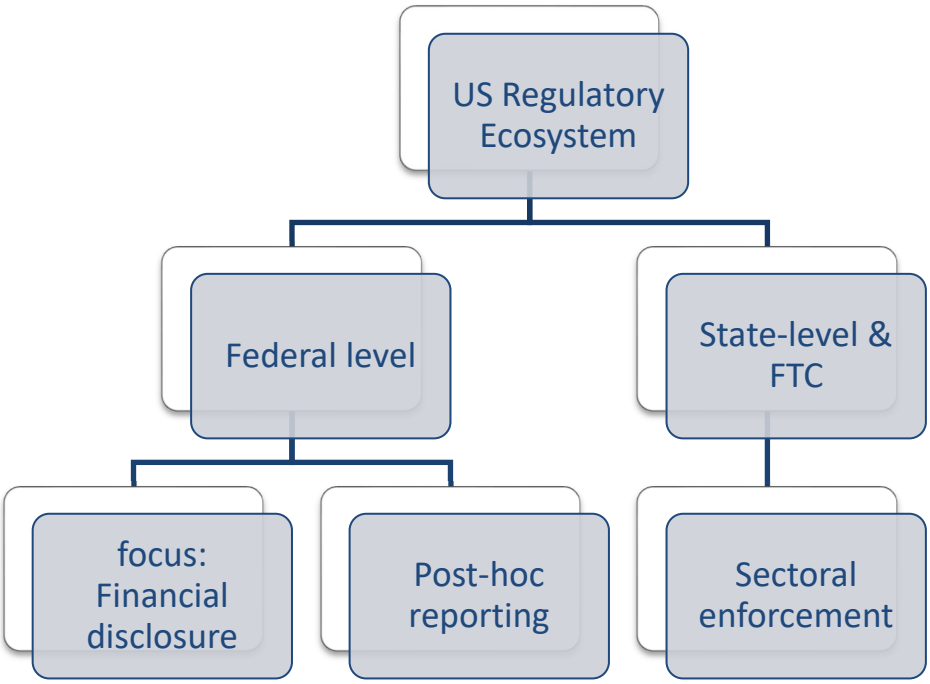


Figure – 1. The United States Framework: Market-Driven Governance and Constitutional Protections

This decentralized approach fosters high levels of innovation and commercial flexibility, allowing tech-driven lobbying startups and corporate entities to rapidly deploy advanced AI tools in Washington. However, it simultaneously creates significant transparency deficits. The reliance on post-hoc disclosure means that manipulative or deceptive algorithmic lobbying campaigns often go undetected until after legislative decisions have already been made. Furthermore, disparities between federal rules and state-level lobbying regulations create a complex compliance mosaic, leaving lawmakers vulnerable to sophisticated digital influence operations that exploit jurisdictional seams [13].

Transatlantic Divergence: Comparative Analysis of Regulatory Efficacy

A rigorous comparative evaluation of the European Union and the United States reveals a fundamental clash between two distinct governance philosophies: preemptive institutional control versus market-driven pluralism. Both jurisdictions grapple with the exact same technological disruption-the integration of artificial intelligence into interest representation – yet their legal responses reflect divergent cultural, political, and constitutional values [14].

The primary divergence lies in how each model handles transparency and accountability. The EU model assumes that algorithmic systems operating within the political sphere pose inherent systemic risks to democracy, thus requiring rigorous, upfront certification, mandatory registration of digital tools, and strict compliance monitoring. This approach effectively minimizes covert corporate manipulation and enhances public visibility into how digital lobbying operates. However, it risks slowing down technological adaptation and creating high administrative barriers for smaller advocacy groups.

Conversely, the US model prioritizes broad protection of political expression and market innovation, relying on transparency-after-the-fact and civil litigation to address abuses. While this preserves constitutional liberties and encourages rapid technological advancement, it leaves democratic institutions acutely vulnerable to sophisticated, opaque algorithmic influence campaigns [15]. The absence of federal constraints on algorithmic lobbying tools means that well-resourced corporations can leverage advanced generative AI to shape public policy with minimal regulatory friction.

Table – 1. Transatlantic Divergence: Comparative Analysis of Regulatory Efficacy

Analytical Category	European Union Model	United States Model
Primary Objective	Democratic integrity, fundamental rights protection, and market harmonization	Free speech protection, market innovation, and minimal state interference
Legislative Style	Comprehensive, centralized, and prescriptive (EU AI Act)	Fragmented, sectoral, and descriptive (LDA, state laws)
Adaptability to AI	High structural adaptability via explicit technology-specific statutes	Low statutory adaptability; relies on retroactive interpretation of legacy laws
Vulnerability Profile	Risk of over-regulation, administrative bottlenecks, and compliance rigidity	Risk of algorithmic opacity, covert corporate capture, and weak oversight

Ultimately, this comparative analysis demonstrates that neither model has achieved an ideal equilibrium. The EU struggles with regulatory overreach and enforcement complexities in a fast-moving technological landscape, while the US faces severe accountability gaps and structural opacity in digital advocacy. As artificial intelligence continues to evolve, bridging these transatlantic gaps will require innovative regulatory mechanisms that balance democratic resilience with technological dynamism.

Conclusion.

The intersection of artificial intelligence and digital lobbying has fundamentally altered the architecture of modern interest representation, bridging the worlds of computational technology and public policy. This comparative analysis of the European Union and the United States reveals that the integration of algorithms, big data analytics, and generative models into advocacy is no longer a futuristic prospect, but a present-day reality reshaping legislative outcomes. However, the ways in which these two legal superpowers govern this technological disruption expose a profound transatlantic divergence rooted in contrasting institutional philosophies, constitutional values, and regulatory traditions.

The European Union has established a proactive, rights-centric paradigm characterized by preemptive compliance, centralized accountability, and robust institutional oversight. Through legislative milestones such as the Artificial Intelligence Act and the stringent enforcement mechanisms of the EU Transparency Register, the Union attempts to immunize its democratic processes against opaque algorithmic manipulation and covert corporate overreach. While this comprehensive model provides high levels of transparency and consumer protection, it simultaneously introduces significant compliance burdens and potential administrative bottlenecks that may challenge smaller civil society actors.

Conversely, the United States operates within a decentralized, market-driven, and litigation-tempered governance model heavily insulated by First Amendment protections. Grounded in legacy statutes like the Lobbying Disclosure Act, the US regulatory apparatus focuses primarily on post-hoc financial disclosure rather than technological restriction. This approach preserves constitutional liberties, fosters rapid technological innovation, and empowers political expression, but it leaves democratic institutions acutely vulnerable to sophisticated, automated influence campaigns and algorithmic opacity. The absence of a unified federal framework specifically addressing digital lobbying creates critical accountability gaps, making it difficult to detect or deter covert corporate manipulation until after legislative decisions are cemented.

Ultimately, neither the EU's prescriptive rigidity nor the US's fragmented permissiveness offers a flawless blueprint for the future of digital governance. As artificial intelligence systems

grow increasingly autonomous, persuasive, and pervasive, policymakers across both jurisdictions must evolve beyond binary choices of heavy-handed restriction or laissez-faire neglect. Ensuring democratic resilience in the age of automated interest intermediation requires dynamic regulatory frameworks that harmonize technological innovation with uncompromising institutional transparency. Only through adaptive legal mechanisms, cross-border regulatory cooperation, and continuous empirical scrutiny can modern democracies safeguard the integrity of public policy against the silent, scalable influence of algorithmic power.

REFERENCES

1. Acemoglu, D., & Johnson, S. (2023). *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*. PublicAffairs.
2. Bertolini, A. (2020). Artificial intelligence and civil liability. *European Parliament Policy Department for Citizens' Rights and Constitutional Affairs*, PE 621.926.
3. Casey, B. C., & Niblett, A. (2020). Self-driving laws. *University of Toronto Law Journal*, 70(4), 429–456.
4. Darrell, M., & West, N. (2022). *The Future of Work: Robotics, AI, and Automation*. Brookings Institution Press.
5. European Commission. (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. COM(2021) 206 final.
6. European Parliament. (2023). *The EU Transparency Register: Strengthening Accountability in Public Decision-Making*. EU Publications Office.
7. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1–15.
8. Gold, M. (2021). Algorithmic lobbying and the transformation of interest intermediation in American politics. *Journal of Public Policy and Governance*, 39(2), 115–130.
9. Habermas, J. (2022). *A New Structural Transformation of the Public Sphere and Deliberative Politics*. Polity Press.
10. Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
11. Ruhl, J. B., & Katz, D. M. (2022). Measuring, monitoring, and managing legal analytics in the age of machine learning. *Northwestern University Law Review*, 116(3), 789–842.
12. Schlozman, K. L., Verba, S., & Brady, H. E. (2020). *The Unheavenly Chorus: Unequal Voice and the Broken Promise of American Democracy*. Princeton University Press.
13. Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, 4(1), 1–45.
14. United States Congress. (1995). *Lobbying Disclosure Act of 1995*. Public Law 104-65, 2 U.S.C. ch. 26.
15. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Political Studies

Cybersecurity as a Factor of Kazakhstan's National Foreign Policy Strategy

Arzigul Islamova
Department of International Relations, Institute Sorbonne-Kazakhstan, Abai Kazakh National Pedagogical University, Almaty, Republic of Kazakhstan
Aimen Kurmangali
Sorbonne-Kazakhstan Institute, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
İhsan Çapcıoğlu
Faculty of Theology, Department of Philosophy and Religious Studies, Division of Sociology of Religion, Ankara University, Turkey

Abstract

This article examines cybersecurity as an increasingly autonomous factor shaping the foreign policy strategy of the Republic of Kazakhstan in the period 2020-2028. Drawing on the Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030, the national "Cyber Shield of Kazakhstan" programme, and the country's evolving engagement with multilateral cyber-diplomacy platforms, the study argues that cybersecurity has moved from a purely technical and domestic-security concern to a structural component of Kazakhstan's multi-vector diplomacy. The authors analyze the institutional architecture responsible for cyber-policy formation, Kazakhstan's positioning between Russia, China, the European Union and the United States on questions of digital sovereignty and information security, and the country's participation in United Nations, OSCE, SCO, CSTO and CIS cyber-related initiatives. Particular attention is paid to the tension between Kazakhstan's aspiration for an open, internationally integrated digital economy and its simultaneous pursuit of information sovereignty and control over the domestic information space. Based on document analysis and comparative case material, the article concludes that cybersecurity functions in contemporary Kazakhstani foreign policy simultaneously as a security imperative, a diplomatic resource, and a legitimizing narrative for the modernization of state institutions. The findings contribute to the broader academic discussion on digital diplomacy and technology-driven transformation of foreign policy doctrines in Central Asian states, and offer a set of policy recommendations aimed at strengthening Kazakhstan's cyber-diplomatic capacity through 2028.

Keywords: *cybersecurity; digital diplomacy; foreign policy; Kazakhstan; multi-vector diplomacy; Cyber Shield of Kazakhstan; information sovereignty; Central Asia; artificial intelligence; international relations*

1. Introduction

Over the past decade, cybersecurity has evolved from a narrow technical discipline concerned with the protection of computer networks into a central pillar of state security strategy and, increasingly, of foreign policy itself. States no longer treat the digital domain as a peripheral extension of domestic administration; rather, they treat it as a contested space in which national interests, alliances, and rivalries are negotiated. For middle powers and post-Soviet states in

particular, cybersecurity has become a field in which international reputation, sovereignty, and strategic autonomy are simultaneously at stake.

The Republic of Kazakhstan offers an instructive case in this respect. Since the adoption of the Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030, approved by presidential decree No. 280 of 6 March 2020, Kazakhstan's foreign policy doctrine has explicitly acknowledged the growing role of information and communication technologies, and of digital security in particular, in shaping the country's international position. This doctrinal shift did not occur in isolation. It followed several years of domestic cybersecurity institution-building, beginning with the 2017 "Cyber Shield of Kazakhstan" concept, and it coincided with the broader "Digital Kazakhstan" state programme aimed at digitizing the economy, public administration, and social services.

This article addresses a specific and, in the authors' view, insufficiently examined question: how has cybersecurity been integrated into Kazakhstan's foreign policy strategy as an autonomous factor, rather than merely as a domestic technical or defensive concern? We argue that cybersecurity in the Kazakhstani case performs at least three distinct functions in foreign policy: (1) a security function, protecting critical infrastructure and state institutions from external and internal threats; (2) a diplomatic function, serving as a platform for cooperation with multiple international partners consistent with the doctrine of multi-vector diplomacy; and (3) a legitimizing function, reinforcing the narrative of Kazakhstan as a modernizing, technologically capable and internationally responsible state.

The significance of this inquiry extends beyond the Kazakhstani case itself. As artificial intelligence and advanced digital technologies increasingly permeate both the instruments of statecraft and the threat landscape that foreign ministries must manage, understanding how a specific state has integrated cybersecurity into its foreign policy doctrine offers a template for analyzing similar processes elsewhere in Central Asia, the wider post-Soviet space, and among middle powers more generally. Kazakhstan is a particularly informative case precisely because it combines relatively advanced digital infrastructure, an explicit and long-standing multi-vector diplomatic tradition, and a geopolitical position between major powers with divergent approaches to cyber-governance, making it a useful laboratory for observing how cybersecurity considerations are refracted through, and in turn reshape, an established foreign policy doctrine.

The analysis proceeds as follows. Section 2 outlines the theoretical framework, situating cybersecurity within the broader literature on digital diplomacy and foreign policy analysis. Section 3 reviews the existing scholarship on Kazakhstan's digital transformation and cyber-policy. Section 4 describes the methodology and sources used. Section 5 examines the place of cybersecurity within the 2020-2030 Foreign Policy Concept. Section 6 analyzes the institutional and legal architecture of Kazakhstan's cybersecurity policy. Section 7 considers cybersecurity as an instrument of multi-vector diplomacy vis-a-vis Russia, China, the European Union, the United States, and multilateral institutions. Section 8 offers an empirical illustration through selected cases. Section 9 discusses risks and structural contradictions. Section 10 presents policy recommendations, and Section 11 concludes.

2. Theoretical Framework: Cybersecurity as a Foreign Policy Variable

The theoretical starting point of this article is the observation that cybersecurity can no longer be adequately understood through a purely technical or purely military lens. Scholars of digital diplomacy, notably Bjola and Holmes, have argued that digital technologies have transformed not only the tools available to diplomats but the very structure of diplomatic practice, creating new channels of representation, negotiation, and public engagement. In parallel, security studies scholars such as Kello have described cyberspace as a domain in which the traditional Clausewitzian logic of war and peace becomes blurred, producing a persistent condition of low-intensity conflict that states must manage diplomatically as much as militarily.

Joseph Nye's concept of soft power provides a further useful lens: a state's capacity to shape outcomes through attraction rather than coercion increasingly depends on its digital infrastructure, its cyber-resilience, and its credibility as a partner in information security cooperation. A state perceived as cyber-vulnerable risks being excluded from sensitive cooperation frameworks, while a state that can demonstrate credible cyber-governance gains bargaining leverage in negotiations over data flows, critical infrastructure investment, and technology transfer.

Comparative cases from outside Central Asia further illuminate the theoretical stakes. Estonia's experience following the 2007 cyberattacks against its government and banking infrastructure is frequently cited as the paradigmatic instance of a small state converting a cyber-crisis into a durable diplomatic and institutional asset, subsequently hosting the NATO Cooperative Cyber Defence Centre of Excellence and building an international reputation as a cyber-governance leader disproportionate to its size. Singapore offers a related but distinct model, in which a technocratic, tightly coordinated cybersecurity agency has been deliberately used as a platform for regional capacity-building across Southeast Asia, reinforcing the city-state's broader diplomatic identity as a reliable, rules-based partner. Both cases illustrate the general proposition, central to this article, that cybersecurity capability and cyber-diplomatic reputation can substitute, in part, for the material power resources that would otherwise be required to achieve comparable international influence, a dynamic of particular relevance to Kazakhstan given its intermediate position between much larger neighbouring powers.

For the purposes of this study, we adopt a three-level analytical framework distinguishing (a) the doctrinal level, comprising official foreign policy concepts and strategic documents; (b) the institutional level, comprising the state bodies, agencies, and legal instruments responsible for cybersecurity governance; and (c) the relational level, comprising bilateral and multilateral diplomatic engagements in which cybersecurity is negotiated, coordinated, or contested. This framework allows us to trace how a technical policy domain is translated into a foreign policy instrument, and conversely how foreign policy priorities shape domestic cybersecurity institution-building.

We further draw on the concept of "digital sovereignty," understood as a state's claim to exercise regulatory and operational control over the data, infrastructure, and information flows located within or connected to its territory. For post-Soviet states such as Kazakhstan, digital sovereignty is closely intertwined with broader questions of political sovereignty and independence from the influence of larger neighbouring powers, which adds a distinctive regional dimension to the general theoretical literature developed primarily with reference to Western and East Asian cases.

A final theoretical element concerns the rapidly growing intersection of cybersecurity and artificial intelligence in international affairs. AI-enabled tools are increasingly used both offensively, for example in automated reconnaissance, disinformation generation, and adaptive malware, and defensively, in threat-detection, anomaly-recognition, and predictive risk-assessment systems deployed by national computer emergency response teams. This convergence has begun to reshape the vocabulary of international cyber-diplomacy, as multilateral discussions on responsible state behaviour in cyberspace increasingly overlap with parallel discussions on the governance of artificial intelligence, including at United Nations bodies and regional organizations in which Kazakhstan participates. For a state such as Kazakhstan, positioning itself early in these overlapping normative conversations offers an opportunity to shape emerging rules rather than merely adapt to them, reinforcing the argument that cybersecurity, augmented by AI considerations, functions as a proactive rather than purely reactive element of foreign policy strategy.

The three-level framework introduced above therefore requires a fourth, cross-cutting dimension: the technological dimension, capturing how the accelerating pace of technological change, and of artificial intelligence in particular, continuously reshapes the doctrinal, institutional, and relational levels of cyber-diplomacy simultaneously, rather than affecting them sequentially. This dimension is particularly salient for the 2020-2028 period under review, which spans both the initial consolidation of Kazakhstan's cybersecurity institutions and the subsequent, much more rapid diffusion of generative and analytical AI tools across government, industry, and society.

3. Literature Review

The academic literature on Kazakhstan's digital transformation has grown considerably since the launch of the "Digital Kazakhstan" state programme in 2017-2018. A first strand of scholarship focuses on the domestic dimension: e-government development, digitization of public services, and the modernization of state administration. This literature generally treats cybersecurity as a supporting, technical component of digital governance rather than as an object of foreign policy analysis in its own right.

A second strand, closer to the concerns of this article, examines Kazakhstan's foreign policy doctrine and its evolution from the early independence period through the 2020-2030 Concept. Scholars such as those writing in the Institute of World Economy and International Relations journal have traced the continuity of Kazakhstan's multi-vector, balanced foreign policy across successive doctrinal iterations, while noting the growing salience of new, non-traditional security domains, including cyber and information security, within these documents.

A third, more specialized strand addresses cybersecurity policy specifically. Analysts associated with the Central Asia security studies community have documented the emergence of the "Cyber Shield of Kazakhstan" concept in 2017, its subsequent action plans, and the country's improvement in international cybersecurity rankings, including the International Telecommunication Union's Global Cybersecurity Index, where Kazakhstan has moved from a relatively low ranking toward the upper tier of participating states. However, this literature has tended to treat cybersecurity primarily as a matter of domestic institution-building and critical infrastructure protection, with comparatively limited attention to its diplomatic and foreign-policy dimensions.

A fourth strand, situated in the broader international relations literature on digital diplomacy, has examined how middle powers and smaller states use technology cooperation, including cybersecurity assistance and capacity-building, as a form of niche diplomacy that allows them to gain visibility and influence disproportionate to their material power. This literature has been applied extensively to states such as Estonia and Singapore, but only sparingly to Central Asian cases, leaving a gap that this article seeks to address.

Taken together, the existing literature provides a solid empirical foundation regarding Kazakhstan's domestic cyber-institutional architecture and its general foreign policy doctrine, but it has not yet systematically connected the two by asking how cybersecurity functions as an instrument and object of foreign policy strategy in its own right. This article aims to fill that gap.

A fifth and more recent strand of scholarship, still emerging at the time of writing, addresses the relationship between artificial intelligence and foreign policy formation in Central Asia more broadly. This literature has so far focused primarily on AI's role in economic diversification strategies and on the risks of algorithmic disinformation in domestic political contexts, with comparatively little attention paid to the specific ways in which AI-enabled capabilities are altering the practice of cyber-diplomacy, threat-intelligence sharing, and multilateral norm negotiation. The present article contributes to this emerging strand by explicitly integrating AI-related developments into its analysis of Kazakhstan's cybersecurity-foreign policy nexus, particularly in Section 9 below.

A sixth strand of relevance, though focused on a neighbouring rather than the Kazakhstani case, examines Uzbekistan's parallel digital and cybersecurity reform trajectory since 2016, including the establishment of its own information-security service and the adoption of comparable data-localization and critical-infrastructure protection measures. This comparative literature is useful for the present study because it highlights both the common regional pressures, such as the shared legacy of Soviet-era telecommunications infrastructure and comparable exposure to cyber threats emanating from outside the region, and the distinct national trajectories that have emerged, with Kazakhstan generally pursuing a more internationally networked and multi-vector approach to cyber-cooperation than several of its immediate neighbours, consistent with its longer-standing and more institutionally developed multi-vector diplomatic tradition.

It should also be noted that much of the available literature relies heavily on official government sources and statements by Kazakhstani officials at international forums, which, while valuable for reconstructing the country's declared doctrine and self-presentation, offer limited independent verification of operational capabilities or the internal decision-making processes underlying cyber-diplomatic initiatives. This article shares this limitation, common to research on relatively closed security-sector institutions, and addresses it by triangulating official statements with independent reporting and comparative regional analysis wherever possible, while explicitly flagging areas where further empirical research, including interview-based fieldwork, would be valuable.

4. Methodology and Sources

This study employs a qualitative, document-based research design combined with comparative case illustration. The primary sources include the Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030 (Decree of the President No. 280, 6 March 2020, with subsequent amendments), the "Cyber Shield of Kazakhstan" cybersecurity concept and its implementation action plans, publicly available materials from the Ministry of Digital Development, Innovations and Aerospace Industry, the Committee for National Security, and statements by Kazakhstani officials at international forums such as the Eurasian Economic Forum and sessions of the United Nations Open-Ended Working Group on developments in the field of information and telecommunications.

Secondary sources include academic and policy analyses of Kazakhstan's digital transformation and foreign policy doctrine published in peer-reviewed journals and by recognized research institutes specializing in Central Asian and Eurasian affairs, as well as reporting by specialized outlets tracking cybersecurity developments in the post-Soviet space.

The analysis proceeds through qualitative content analysis of doctrinal documents, identifying explicit and implicit references to cybersecurity, information security, digital sovereignty, and technology cooperation, and situating these references within the broader structure of the foreign policy concept. This is complemented by an institutional mapping exercise identifying the principal state bodies responsible for cyber-policy and their international counterparts, and by a comparative review of Kazakhstan's cyber-related engagements with Russia, China, the European Union, the United States, and relevant multilateral bodies.

The study period, 2020-2028, was chosen to correspond with the adoption of the current Foreign Policy Concept in 2020 and to extend analytically to the horizon envisaged by that document, allowing the article to combine retrospective document analysis for the period already elapsed at the time of writing with a forward-looking, scenario-informed discussion, particularly in Section 9, of likely developments through the remainder of the decade. Where the analysis addresses developments beyond the immediate historical record, this is explicitly flagged as an informed projection rather than a description of settled fact, consistent with the qualitative and interpretive character of the study as a whole.

Given the qualitative and interpretive nature of the analysis, the article does not claim statistical generalizability; rather, it aims to provide an analytically grounded account of the mechanisms through which cybersecurity has become integrated into Kazakhstan's foreign policy strategy, which may serve as a basis for further comparative research on other Central Asian states.

5. Cybersecurity in the Foreign Policy Concept of Kazakhstan for 2020-2030

The Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030 explicitly situates the country's diplomacy within a period of complex transformation of the international system, characterized by a crisis of confidence among major powers, erosion of multilateral security mechanisms, and the emergence of new, non-traditional security threats. Among these threats, the document identifies the growing significance of information and communication technologies, both as instruments of development and as sources of new vulnerabilities, including cyberattacks and hybrid, "new type" forms of confrontation.

The Concept reaffirms the continuity of Kazakhstan's multi-vector, pragmatic, and proactive foreign policy, seeking to balance relationships with Russia and China as immediate strategic partners, the United States and the European Union as leading global powers, and the states of Central Asia as Kazakhstan's closest neighbourhood, where the country has increasingly positioned itself as a regional leader. Cybersecurity cooperation fits naturally within this multi-vector logic: it allows Kazakhstan to engage simultaneously with partners whose broader strategic interests may diverge, without requiring an exclusive geopolitical alignment.

The doctrine's emphasis on digital transformation as a driver of economic modernization further reinforces the link between cybersecurity and foreign policy. Kazakhstan's aspiration to attract foreign investment, participate in international value chains, and position itself as a regional digital and logistics hub within the Digital Silk Road and similar connectivity initiatives depends on its ability to demonstrate credible standards of data protection and critical infrastructure security. In this sense, cybersecurity functions doctrinally not only as a defensive concern but as an enabling condition for the achievement of broader economic and diplomatic objectives articulated in the 2020-2030 Concept.

Amendments introduced to the Concept in subsequent years, including targeted revisions announced in 2024-2025, have continued to reflect the increasingly demanding character of multi-vector diplomacy amid heightened geopolitical turbulence, reinforcing rather than diminishing the salience of technological and information-security considerations in the document's underlying logic.

It is also significant that the 2020-2030 Concept was adopted at the outset of President Kassym-Jomart Tokayev's tenure, in the context of a broader domestic reform agenda emphasizing the modernization of state institutions and a more transparent, "listening state" model of governance. Cybersecurity and digital transformation fit naturally within this narrative: a state that can present itself as digitally capable and cyber-resilient reinforces its broader claim to administrative modernization, which in turn supports its diplomatic messaging toward both Western partners, who tend to value institutional modernization as a marker of reform, and Eastern partners, who tend to value demonstrated technical and administrative capacity as a marker of reliability.

The Concept's treatment of regional leadership in Central Asia further illustrates the link between cybersecurity and doctrine. As Kazakhstan has sought to consolidate its position as the leading state in Central Asia, extending cooperation to Uzbekistan, Kyrgyzstan, Tajikistan, and Turkmenistan across a range of sectors, cybersecurity and digital connectivity have emerged as natural areas for such regional leadership, given the relatively advanced state of Kazakhstan's digital infrastructure compared to some of its immediate neighbours. Kazakhstan's periodic offers to share cybersecurity tools and expertise with regional partners, discussed further in Section 8,

should therefore be read not only as bilateral gestures but as expressions of the doctrinal ambition, explicit in the 2020-2030 Concept, to secure the status of a leading regional state.

The Concept also situates cybersecurity within Kazakhstan's economic diplomacy toward international financial and development institutions. As Kazakhstan has sought continued engagement with the World Bank, the Asian Development Bank, and the European Bank for Reconstruction and Development on digital-infrastructure financing, demonstrable progress on cybersecurity and data-protection standards has functioned as a precondition for certain categories of investment, particularly in sectors such as digital financial services and e-government platforms where donor institutions apply their own due-diligence requirements regarding data security and system resilience. This dimension situates cybersecurity not only within the security and political vectors discussed above but also within Kazakhstan's development-finance diplomacy, a linkage that is likely to become more pronounced as digital-infrastructure investment needs grow through 2028.

Finally, the Concept's framing of cybersecurity sits within a broader securitization of the information domain that extends to concerns about disinformation, foreign interference in domestic political processes, and the protection of what Kazakhstani officials frequently describe as the country's "information sovereignty." This framing links cybersecurity policy to a wider set of concerns about media regulation and public communication that, while not the primary focus of this article, form an important part of the domestic political context within which Kazakhstan's international cyber-diplomacy is formulated and legitimized.

6. Institutional and Legal Architecture of Kazakhstan's Cybersecurity Policy

The institutional foundation of Kazakhstan's cybersecurity policy was laid with the adoption, in 2017, of the "Cyber Shield of Kazakhstan" concept, developed in the context of the presidential programme "The Third Modernization of Kazakhstan: Global Competitiveness" and the long-term development strategy "Kazakhstan-2050." The concept was explicitly framed as Kazakhstan's national contribution to the implementation of the United Nations' global cybersecurity agenda, signalling from the outset an intention to link domestic cyber-institution-building with the country's international standing.

The action plan accompanying the Cyber Shield concept, approved in October 2017, tasked a range of state bodies, including the Committee for National Security, the (then) Ministry of Defense and Aerospace Industry, and later the Ministry of Digital Development, Innovations and Aerospace Industry, with developing a modern and reliable system for the prevention and mitigation of cyberattacks and hybrid threats. The plan reflected concerns voiced at the highest political level: President Nursultan Nazarbayev instructed the government to treat cybersecurity as a matter of national security, observing that in the contemporary world critical infrastructure, such as power plants or railway networks, could be paralyzed without recourse to conventional military means.

Implementation of the Cyber Shield concept has involved multiple stakeholders beyond the security services, including the banking sector, the national telecommunications operators Kazakhtelecom and Transtelecom, the Centre for Analysis and Investigation of Cyberattacks, the International University of Information Technologies, and industry associations such as the Kazakhstan Information Security Association. This multi-stakeholder structure has allowed Kazakhstan to combine state-led coordination with technical expertise drawn from the private sector and academia, a model that officials have subsequently presented internationally as a template for cooperation with partner states.

By the mid-2020s, officials reported that these efforts had contributed to a substantial improvement in Kazakhstan's position in the International Telecommunication Union's Global Cybersecurity Index, with the Vice Minister of Digital Development citing a rise from around the 108th position to the 31st position among ranked countries, alongside the development of a

package of instruments for the protection of state and personal data. This improvement has itself become a diplomatic asset: Kazakhstani representatives have used it to offer cybersecurity cooperation, including cryptographic tools for the protection of cross-border data transfers, to partner states within the Eurasian Economic Union and beyond.

On the legal side, Kazakhstan's cybersecurity architecture rests on legislation governing informatization, personal data protection, and critical information infrastructure, complemented by government resolutions establishing computer emergency response functions and information security requirements for state and quasi-state organizations. While a comprehensive review of this legislative corpus lies beyond the scope of the present article, its overall direction has been toward greater centralization of cyber-governance under the national security apparatus, combined with sector-specific requirements for the financial, telecommunications, and energy industries deemed critical infrastructure.

The framework law "On Informatization," together with subsequent amendments addressing personal data protection and the licensing of information-security activities, establishes baseline obligations for both public bodies and private operators of information systems classified as critical. These obligations include mandatory security certification of software and hardware used in critical infrastructure, incident-reporting requirements, and, in certain sectors, data-localization provisions requiring that specified categories of data be stored on servers located within Kazakhstani territory. Such data-localization requirements are directly relevant to foreign policy, since they shape the terms on which foreign technology companies, cloud-service providers, and financial institutions may operate in Kazakhstan, and they have at times required diplomatic clarification in the context of Kazakhstan's commitments under Eurasian Economic Union integration and its broader trade and investment relationships with the European Union and the United States.

Kazakhstan's legal framework also increasingly references cooperation with international standard-setting bodies, including alignment, in certain respects, with international information-security management standards such as ISO/IEC 27001, which domestic regulators have encouraged critical-infrastructure operators to adopt as a means of demonstrating compliance with internationally recognized best practice. This orientation toward international standards, rather than purely domestic or bloc-specific technical requirements, provides a further mechanism through which Kazakhstan's legal architecture supports, rather than constrains, the multi-vector diplomatic posture examined in Section 7.

This institutional configuration has direct implications for foreign policy. Because cybersecurity governance in Kazakhstan is closely tied to the national security apparatus rather than being distributed primarily among civilian digital-economy regulators, cybersecurity cooperation with foreign partners frequently takes place through security and intelligence channels in parallel with, or in place of, purely technical or commercial channels. This has consequences for the character of Kazakhstan's cyber-diplomacy, which we examine in the following section.

A further institutional development of note is the establishment and gradual strengthening of a national computer emergency response capability, tasked with coordinating incident response, vulnerability disclosure, and threat-intelligence sharing both domestically and with foreign counterpart agencies. Participation in international and regional computer emergency response team networks provides Kazakhstan with an additional, more technically oriented diplomatic channel, distinct from the primarily security-driven channels described above, through which the country can build professional relationships with counterpart agencies in Europe, Asia, and the Commonwealth of Independent States without the same degree of political sensitivity that attaches to higher-level security cooperation.

The financial sector deserves particular mention within this institutional architecture, given its systemic importance and its frequent exposure to cyberattacks, as noted in early Cyber Shield documentation describing the vulnerability of domestic financial institutions. The National Bank of the Republic of Kazakhstan has been an active participant in the Cyber Shield implementation process, reflecting the broader international trend of treating financial-sector cybersecurity as a distinct sub-domain requiring specialized regulatory attention, technical standards, and often dedicated international cooperation channels, including with international financial institutions and foreign central banks.

Coordination between civilian and security-oriented cyber-institutions has historically posed challenges common to many states undergoing rapid digital-institutional expansion. Overlapping mandates between the Ministry of Digital Development and the Committee for National Security, for instance, have at times required clarifying government resolutions to delineate responsibility for licensing, certification, and incident-response functions. From a foreign policy perspective, such internal coordination challenges matter because foreign partners negotiating cybersecurity cooperation agreements require clarity regarding which Kazakhstani institution holds effective authority over a given aspect of implementation, and unresolved overlaps can slow the conclusion or execution of international agreements, particularly those involving sensitive data-sharing or joint incident-response arrangements.

Educational and human-capital institutions, including the International University of Information Technologies and specialized programmes at other national universities, form a further, less frequently discussed component of the institutional architecture. Their inclusion in the Cyber Shield implementation structure reflects an official recognition that sustainable cybersecurity capacity requires a domestic talent pipeline, and their growing international academic partnerships, including with European and Asian universities, constitute an additional, quieter channel of cyber-diplomacy operating alongside the more visible state-to-state engagements analyzed elsewhere in this article.

7. Cybersecurity as an Instrument of Multi-Vector Diplomacy

Kazakhstan's foreign policy doctrine of multi-vector diplomacy, reaffirmed in the 2020-2030 Concept, commits the country to balanced engagement with Russia and China as immediate strategic partners, with the United States and the European Union as leading global powers, and with multilateral institutions as platforms for the pursuit of national interests. Cybersecurity cooperation reproduces this multi-vector logic across at least four distinct relational contexts.

7.1. Cooperation with Russia and Eurasian institutions. Kazakhstan's cybersecurity cooperation with Russia is conducted principally through the frameworks of the Collective Security Treaty Organization and the Commonwealth of Independent States, both of which have developed information-security cooperation mechanisms addressing cross-border cybercrime, critical infrastructure protection, and coordinated responses to information threats. Kazakhstan has also offered to extend its own cybersecurity and cryptographic tools to partners within the Eurasian Economic Union, reflecting an ambition to position itself not merely as a recipient but as a provider of regional cybersecurity capacity.

In evaluating the Russian vector further, it is worth noting that Kazakhstan's cooperation with Russia on cybersecurity operates against the backdrop of deep pre-existing integration in telecommunications infrastructure, payment systems, and technical standards inherited from the Soviet period and reinforced through subsequent Eurasian integration. This legacy infrastructure interdependence gives cybersecurity cooperation with Russia a qualitatively different character from Kazakhstan's cooperation with other partners: it is less a matter of building new cooperative mechanisms from scratch and more a matter of managing, updating, and periodically renegotiating an already dense web of technical and institutional linkages, a dynamic that

constrains the pace at which Kazakhstan can diversify away from Russian-linked infrastructure even where it may wish to do so for reasons of risk diversification.

7.2. Cooperation with China. Engagement with China occurs primarily through the Shanghai Cooperation Organization's information-security track, which frames cyber and information threats in terms closely aligned with concepts of information sovereignty and non-interference, as well as through bilateral cooperation associated with digital infrastructure development linked to the Belt and Road Initiative and its "Digital Silk Road" component. This cooperation offers Kazakhstan access to infrastructure investment and technology, but also raises questions, discussed further in Section 9, about technological dependency and the compatibility of imported governance models with Kazakhstan's stated aspiration for an open digital economy.

The Chinese vector also illustrates a broader pattern in which infrastructure investment and cybersecurity governance are bundled together in practice, even where they are formally negotiated through separate channels. Digital-infrastructure projects associated with the Digital Silk Road, including fibre-optic backbone links, data centres, and smart-city applications, typically incorporate security and surveillance components supplied by the same firms providing the underlying hardware, meaning that decisions about infrastructure procurement carry embedded cybersecurity-governance implications that are not always explicitly negotiated as such, a pattern that has attracted growing scrutiny from both Kazakhstani civil-society observers and Western commentators concerned about technological dependency.

7.3. Cooperation with the United States and the European Union. Kazakhstan's engagement with Western partners on cybersecurity has taken a comparatively more technical and capacity-building character, including cooperation with private international cybersecurity firms and participation in cyber-capacity-building initiatives supported by Western governments and organizations such as the OSCE. Kazakhstan's chairmanship of the OSCE in 2010 and its continued active participation in the organization have provided a platform for engagement with the OSCE's confidence-building measures on the use of information and communication technologies, which seek to reduce the risk of conflict escalation arising from cyber incidents between participating states.

7.4. Multilateral engagement. At the United Nations, Kazakhstan has participated in the Open-Ended Working Group and related processes addressing state behaviour in cyberspace, generally supporting the elaboration of universally applicable norms of responsible state conduct while emphasizing the principle of state sovereignty over national information space, a position broadly consistent with that of Russia and China but not identical to it, reflecting Kazakhstan's characteristic effort to avoid full alignment with any single bloc.

Across these four vectors, cybersecurity functions less as a fixed policy position and more as a flexible diplomatic resource that can be deployed differently depending on the partner, allowing Kazakhstan to extract technology, investment, training, and international legitimacy from multiple directions simultaneously, consistent with the broader multi-vector logic that has characterized Kazakhstani diplomacy since independence.

7.5. Regional leadership within Central Asia. In addition to the four vectors outlined above, Kazakhstan has increasingly used cybersecurity and digital-infrastructure cooperation as an instrument of regional leadership within Central Asia itself. Bilateral consultations and memoranda of understanding with Uzbekistan, Kyrgyzstan, and other neighbouring states on information security, digital-economy standards, and cross-border data protection complement Kazakhstan's participation in regional security structures, including the Regional Anti-Terrorist Structure of the Shanghai Cooperation Organization, which addresses cyber-enabled extremist propaganda alongside conventional counter-terrorism concerns. This regional vector allows Kazakhstan to position itself simultaneously as a recipient of cybersecurity cooperation from larger powers and as a provider of cybersecurity capacity to smaller regional neighbours, reinforcing the doctrinal

ambition, articulated in the 2020-2030 Concept, of consolidating Kazakhstan's status as the leading state in Central Asia.

7.6. Kazakhstan's positioning between competing normative frameworks deserves further elaboration. Western partners, including the European Union and the United States, have generally promoted a normative framework emphasizing a free, open, and interoperable internet, subject to agreed rules against attacks on critical infrastructure and against interference in electoral processes. Russia and China, by contrast, have promoted a framework emphasizing state sovereignty over national information space and a more restrictive interpretation of permissible content and cross-border information flows. Kazakhstan's diplomatic statements in multilateral forums have generally sought a middle position, endorsing the applicability of existing international law to cyberspace and supporting confidence-building measures associated with the Western-led approach, while simultaneously endorsing sovereignty-centred language associated with the Russian- and Chinese-led approach, particularly regarding the protection of critical information infrastructure and the prevention of what Kazakhstani statements describe as the use of information and communication technologies for purposes incompatible with international peace and security. This deliberately ambiguous positioning should be read as a diplomatic strategy in its own right, consistent with the broader multi-vector tradition, rather than as an indication of doctrinal incoherence.

7.7. The role of international organizations beyond the United Nations. Beyond the UN processes discussed above, Kazakhstan's engagement with the International Telecommunication Union, whose Global Cybersecurity Index has served as an important external benchmark and diplomatic reference point for Kazakhstani officials, illustrates a further, more technocratic vector of cyber-diplomacy. By actively working to improve its ranking in such indices, Kazakhstan pursues a form of "indicator diplomacy," using internationally recognized metrics to validate domestic reform narratives and to strengthen its credibility in bilateral and multilateral negotiations on technology cooperation, investment, and data-sharing arrangements.

Taken together, these six vectors demonstrate that Kazakhstan's cyber-diplomacy is not reducible to a single bilateral relationship or a single normative orientation. Rather, it constitutes a deliberately diversified portfolio of engagements, each serving distinct but complementary functions within the country's overarching multi-vector foreign policy strategy, and each carrying its own balance of opportunities and constraints, as discussed further in Section 9.

8. Empirical Illustration: Selected Cases

To illustrate the mechanisms described above, this section briefly considers three representative episodes in Kazakhstan's recent cyber-diplomatic practice.

Before turning to the cases themselves, it is useful to recall the analytical criteria used to select them. Each case was chosen to illustrate a distinct mechanism identified in the preceding theoretical and institutional discussion: the conversion of domestic achievement into diplomatic offer, the blending of commercial and state-security registers in technical cooperation, the strategic ambiguity pursued in multilateral normative engagement, and the use of cybersecurity cooperation as an instrument of regional leadership. Read together, the four cases are intended to demonstrate the range of registers, security, commercial, normative, and regional, through which cybersecurity operates as a foreign policy instrument in the Kazakhstani case, rather than to provide an exhaustive catalogue of all relevant episodes.

The first concerns Kazakhstan's presentation of its cybersecurity achievements at regional economic forums. At the first Eurasian Economic Forum, held in Bishkek, Kazakhstan's Vice Minister of Digital Development, Innovations and Aerospace Industry publicly presented the country's progress under the Cyber Shield programme, including its improved international cybersecurity ranking, and explicitly offered to share cybersecurity instruments, including cryptographic solutions for secure cross-border data transfer, with other Eurasian Economic Union

member states. This episode exemplifies how a domestic cybersecurity achievement is actively converted into a diplomatic offer, reinforcing Kazakhstan's image as a regional technological leader within the Eurasian integration project.

The second concerns technical cooperation agreements between Kazakhstani and foreign cybersecurity firms operating in support of the Cyber Shield concept, such as the cooperation established between international vendors and domestic firms to integrate threat-monitoring and intrusion-prevention solutions into Kazakhstan's information security ecosystem. Such agreements, while commercial in form, are frequently framed publicly as contributions to a state-led national security programme, illustrating the blurring of commercial, technical, and diplomatic registers that characterizes Kazakhstan's approach to cybersecurity cooperation.

The third concerns Kazakhstan's participation in multilateral norm-development processes on responsible state behaviour in cyberspace, where Kazakhstani representatives have consistently paired support for cooperative, rules-based approaches to cybersecurity with an emphasis on the sovereign right of states to regulate their own information space. This dual positioning allows Kazakhstan to engage constructively with both Western-led and Russian- and Chinese-led normative frameworks without fully committing to either, a pattern consistent with the multi-vector logic identified throughout this article.

A fourth case concerns Kazakhstan's regional cybersecurity cooperation within Central Asia, illustrating the fifth vector identified in Section 7. Bilateral consultations between Kazakhstani and Uzbekistani digital-development and security agencies, addressing cross-border data protection, joint incident-response coordination, and harmonization of information-security standards, reflect a deliberate effort to extend Kazakhstan's cybersecurity model to its immediate neighbourhood. Similar, though generally less institutionalized, engagement with Kyrgyzstan and Tajikistan reinforces Kazakhstan's self-presentation as the natural coordinating hub for cybersecurity cooperation within Central Asia, complementing its participation in the Shanghai Cooperation Organization's Regional Anti-Terrorist Structure, which addresses the cyber-enabled dimension of extremist propaganda and recruitment across the region.

Taken together, these cases suggest that Kazakhstan's cyber-diplomacy operates through a recurring pattern: domestic institutional achievements and legislative developments are translated into diplomatic capital, which is then deployed flexibly across multiple bilateral and multilateral channels in a manner consistent with the country's broader multi-vector foreign policy strategy.

9. Artificial Intelligence and the Evolving Cyber-Diplomatic Agenda (2025-2028)

The convergence of artificial intelligence and cybersecurity constitutes one of the most consequential developments likely to shape Kazakhstan's foreign policy strategy through the remainder of the period under review. Three dimensions of this convergence deserve particular attention.

The first is defensive: the integration of AI-based anomaly-detection and predictive-analytics tools into national computer emergency response and critical-infrastructure monitoring systems, which Kazakhstani agencies, following a pattern common among Global Cybersecurity Index leaders, have begun to explore as a means of compensating for persistent shortages of trained cybersecurity personnel. Diplomatically, this creates new opportunities for technology-transfer partnerships with states and firms possessing advanced AI-security capabilities, adding a further layer to the multi-vector portfolio described in Section 7.

The second is offensive and adversarial: the growing use of AI-enabled tools for automated reconnaissance, credential-harvesting, and, of particular relevance to foreign policy, the generation of large-scale disinformation campaigns capable of targeting Kazakhstan's domestic information space or its diplomatic messaging abroad. Kazakhstani officials have increasingly framed such threats as part of a broader hybrid-warfare landscape, echoing language used since

the original 2017 Cyber Shield concept, but now applied to a qualitatively new generation of AI-generated synthetic content, including deepfakes and automated influence operations, that existing legal and institutional frameworks were not originally designed to address.

The third dimension concerns the normative and diplomatic level: the gradual merging of international discussions on responsible state behaviour in cyberspace with parallel discussions on the governance of artificial intelligence, taking place across UN bodies, regional organizations, and technical standard-setting forums. Kazakhstan's participation in these overlapping conversations, while still at an early stage, represents an opportunity to shape emerging international norms on military and dual-use applications of AI in ways consistent with its declared preference for multilateral, sovereignty-respecting frameworks, while also requiring a level of specialized diplomatic and technical expertise that, as noted in Section 3, remains under-developed relative to the pace of technological change.

For Kazakhstan's foreign policy doctrine through 2028, the practical implication is that cybersecurity can no longer be treated as a bounded, discrete policy domain; it is increasingly inseparable from broader questions of AI governance, technological sovereignty, and strategic technology partnerships, a fact that will likely require further doctrinal elaboration beyond the general references contained in the current 2020-2030 Foreign Policy Concept.

10. Risks, Constraints, and Structural Contradictions

Notwithstanding the diplomatic opportunities described above, the integration of cybersecurity into Kazakhstan's foreign policy strategy is subject to several structural tensions that merit critical attention.

First, there is a tension between Kazakhstan's stated ambition to build an open, internationally competitive digital economy, capable of attracting foreign technology investment, and its parallel pursuit of information sovereignty and centralized control over the domestic information space. Excessive reliance on any single external partner's technology, whether Russian, Chinese, or Western, risks constraining Kazakhstan's ability to pursue an independent multi-vector cyber-policy and may expose the country's critical infrastructure to dependency-related vulnerabilities that are difficult to mitigate through diplomatic means alone.

Second, the institutional centralization of cybersecurity governance within the national security apparatus, while offering coordination advantages, may limit the participation of civilian regulators, the private sector, and academic institutions in shaping cyber-diplomatic strategy, potentially narrowing the range of expertise available to policymakers and reducing transparency in decisions affecting the broader digital economy.

Third, Kazakhstan's growing capacity in cyber-defence and information-security cooperation coexists with persistent capacity gaps, including a shortage of qualified cybersecurity professionals, uneven digital literacy across the population, and continuing dependence on foreign vendors for critical hardware and software components, all of which constrain the country's ability to project cyber-capability independently in its diplomatic engagements.

It is also worth noting that the tension between openness and control described above is not static but has shifted over the study period. Domestic legislative and regulatory developments after 2020 have, at various points, tightened requirements on foreign digital-platform operators and expanded the categories of content subject to removal or restriction on national-security or information-security grounds. Whatever the domestic policy rationale for such measures, they carry diplomatic consequences: they periodically become subjects of concern in Kazakhstan's dialogue with European institutions and civil-society organizations monitoring digital rights, requiring the Ministry of Foreign Affairs to manage a recurring diplomatic balancing act between defending the country's sovereign regulatory choices and preserving its broader reputation as a reform-oriented, internationally engaged state.

Fourth, the framing of information security in terms that emphasize state control over the domestic information space raises questions, familiar from the broader literature on digital authoritarianism, about the compatibility of Kazakhstan's approach with international norms on freedom of expression and access to information, a tension that partner states and international organizations are likely to raise with increasing frequency as Kazakhstan's international cyber-diplomatic profile grows.

Fifth, the rapid integration of artificial intelligence tools into both diplomatic practice and cybersecurity operations introduces additional risks, including the potential for AI-enabled disinformation campaigns, adversarial manipulation of automated threat-detection systems, and an evolving requirement for specialized diplomatic and technical expertise capable of addressing AI-cybersecurity interactions at the multilateral level, a domain in which Kazakhstan's institutional capacity remains at an early stage of development.

Sixth, Kazakhstan's geographic and economic exposure to regional instability, including spillover effects from conflicts and sanctions regimes affecting neighbouring states, creates additional cyber-related risks that are not always adequately captured by conventional cybersecurity planning. Sanctions-driven rerouting of trade, finance, and technology flows through Kazakhstan has, on occasion, increased scrutiny of the country's compliance and cybersecurity practices by Western regulators and financial institutions, adding a further layer of diplomatic complexity to Kazakhstan's multi-vector positioning, since demonstrating robust cyber and compliance standards becomes simultaneously a defensive necessity and a reputational requirement for maintaining access to Western financial and technology markets.

Seventh, the asymmetry of technical capacity between Kazakhstan and its principal cyber-diplomatic partners means that cooperation agreements, however formally reciprocal, often in practice involve a one-directional transfer of standards, technology, and, at times, oversight expectations from partner to host state. This asymmetry is not unique to Kazakhstan and characterizes much of middle-power cyber-diplomacy globally, but it nonetheless constrains the degree to which Kazakhstan can be said to exercise full autonomy in shaping the terms of its cybersecurity partnerships, even as it benefits materially from them.

Eighth, the pace of technological change itself constitutes a standing risk to the coherence of foreign policy doctrine. Documents such as the 2020-2030 Concept, drafted at a particular technological moment, necessarily lag behind subsequent developments, including the rapid mainstreaming of generative artificial intelligence after 2022-2023, which was not anticipated in comparable detail at the time the Concept was adopted. This creates a structural gap between doctrine and practice that Kazakhstani policymakers must continually manage through subsidiary strategies, action plans, and ad hoc diplomatic initiatives rather than through the foundational doctrinal document itself, a pattern likely to recur with future waves of technological change extending beyond the present study's 2028 horizon.

These tensions do not negate the diplomatic value of cybersecurity cooperation documented in previous sections, but they do suggest that the sustainability of Kazakhstan's cyber-diplomatic strategy through 2028 will depend significantly on the country's ability to manage these structural contradictions rather than simply expanding the scope of its external partnerships.

A further, cross-cutting constraint concerns coordination costs. As the number of parallel cyber-diplomatic vectors identified in Section 7 has grown, so too has the administrative and analytical burden of ensuring that commitments made within one framework, for example technical standards adopted through cooperation with a Western partner, do not conflict with commitments or expectations arising from cooperation with Russia or China. Without a dedicated coordinating function of the kind proposed in Section 11, the multi-vector approach risks

producing fragmented or even contradictory cyber-policy outcomes, undermining rather than reinforcing Kazakhstan's diplomatic flexibility.

11. Policy Recommendations

Based on the preceding analysis, the authors propose the following recommendations for Kazakhstani policymakers responsible for the intersection of cybersecurity and foreign policy through 2028.

First, institutionalize a dedicated cyber-diplomacy function within the Ministry of Foreign Affairs, working in close coordination with the Committee for National Security and the Ministry of Digital Development, in order to ensure that cybersecurity cooperation with foreign partners is coherently integrated with the broader multi-vector foreign policy strategy rather than pursued separately through security or technical channels alone.

Second, diversify technology partnerships across the multi-vector spectrum in a manner that actively reduces critical dependency on any single external provider, including through targeted investment in domestic cybersecurity research, education, and industry capacity, consistent with the objectives already articulated in the Cyber Shield concept.

Third, expand Kazakhstan's role as a provider, rather than solely a recipient, of regional cybersecurity capacity-building, building on the precedent of offering cryptographic and information-security tools to Eurasian Economic Union partners, and extending similar cooperation to Central Asian neighbours as part of Kazakhstan's aspiration to regional leadership articulated in the 2020-2030 Foreign Policy Concept.

Fourth, develop a clearer and more transparent public articulation of Kazakhstan's positions in multilateral cyber-norm negotiations, distinguishing explicitly between legitimate concerns regarding critical infrastructure protection and cybercrime, on the one hand, and questions of content regulation and freedom of expression, on the other, in order to strengthen the credibility of Kazakhstan's engagement with Western partners and international organizations.

Fifth, invest in specialized training for diplomatic personnel at the intersection of artificial intelligence and cybersecurity, given the accelerating relevance of AI-enabled threats and AI-assisted diplomatic practice, and establish dedicated analytical capacity within the Ministry of Foreign Affairs to monitor and anticipate developments in this rapidly evolving domain through the remainder of the 2020-2028 period covered by this study.

Sixth, strengthen inter-agency and public-private coordination mechanisms, building on the existing multi-stakeholder model established under the Cyber Shield concept, to ensure that decisions regarding international cybersecurity cooperation reflect input from the financial sector, telecommunications operators, and academic institutions in addition to the security services.

Seventh, formalize Kazakhstan's participation in overlapping cybersecurity and artificial-intelligence governance discussions at the United Nations and regional organizations through a single, clearly mandated coordinating body, tasked with ensuring consistency across the country's positions in different multilateral forums and with anticipating the diplomatic implications of fast-moving technological developments before they translate into binding international commitments.

Eighth, commission periodic, independently conducted assessments of Kazakhstan's cyber-diplomatic performance, benchmarked against comparable middle powers and neighbouring Central Asian states, in order to provide policymakers with an evidence base for adjusting the balance among the multiple vectors of cooperation described in this article as circumstances evolve through 2028 and beyond.

12. Conclusion

This article has argued that cybersecurity has become a structural, rather than merely incidental, factor in the foreign policy strategy of the Republic of Kazakhstan during the period 2020-2028. Beginning with the 2017 Cyber Shield of Kazakhstan concept and consolidated through

the Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030, cybersecurity has been progressively integrated into Kazakhstan's multi-vector diplomatic strategy, serving simultaneously as a security imperative, a diplomatic resource deployable across relationships with Russia, China, the European Union, the United States, and multilateral institutions, and a legitimizing narrative supporting the broader project of state modernization.

At the same time, the analysis has identified significant structural tensions, including the risk of technological dependency, the centralization of cyber-governance within the security apparatus, persistent capacity constraints, and normative ambiguities regarding information sovereignty and freedom of expression, which will continue to shape the trajectory of Kazakhstan's cyber-diplomacy through the remainder of the decade.

The Kazakhstani case contributes to the broader academic discussion on digital diplomacy by illustrating how a middle power situated between larger regional powers can use cybersecurity cooperation as a flexible instrument of multi-vector diplomacy, extracting technology, investment, and international legitimacy from multiple directions without committing exclusively to any single geopolitical bloc. Further comparative research, extending this framework to other Central Asian states and to the growing role of artificial intelligence in both cyber-operations and diplomatic practice, would help to clarify whether the patterns identified here are specific to Kazakhstan or represent a more general feature of technology-driven foreign policy transformation in the post-Soviet space.

Finally, this study underscores that the analytical separation between "cybersecurity policy" and "foreign policy," still common in much of the specialized literature, is becoming increasingly artificial in practice. For a state pursuing an explicitly multi-vector strategy under conditions of accelerating technological change, cyber-capability and cyber-cooperation are not merely instruments in service of pre-existing foreign policy goals; they actively help define what those goals can realistically be. Understanding Kazakhstan's foreign policy doctrine for 2020-2028, and its likely evolution beyond that horizon, therefore requires treating cybersecurity and, increasingly, artificial intelligence not as peripheral technical annexes to diplomacy but as constitutive elements of the doctrine itself.

References

1. Decree of the President of the Republic of Kazakhstan No. 280 dated 6 March 2020, "On the Concept of Foreign Policy of the Republic of Kazakhstan for 2020-2030." Astana: Akorda, 2020.
2. National Security Committee of the Republic of Kazakhstan. "Cyber Shield of Kazakhstan" Cybersecurity Concept. Official portal of the Government of the Republic of Kazakhstan, gov.kz.
3. Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan. Materials on the "Digital Kazakhstan" State Programme, 2017-2022.
4. Gussarova, A. "Kazakhstan Launches 'Cyber Shield' Concept." Eurasia Daily Monitor, Jamestown Foundation, vol. 14, issue 150, 20 November 2017.
5. Kaspex.kz. "Ministry of Defense and Aerospace Industry of the Republic of Kazakhstan Summed Up the First Results of the Cybershield of Kazakhstan."
6. Kazinform International News Agency. "Kazakhstan Ready to Share Achievements in Cyber Security." Report on the First Eurasian Economic Forum, Bishkek.
7. Times of Central Asia. "Kazakhstan Updates Its Foreign Policy Concept in a More Difficult Era."
8. Institute of World Economy and International Relations (IMEMO RAS). "Diplomacy of the Republic of Kazakhstan: Evolution of Foreign Policy Concepts." Мировая экономика и международные отношения, 2023, vol. 67, no. 7.
9. Ministry of Foreign Affairs of the Republic of Kazakhstan. "MFA Collegium: Kazakhstan's Multi-Vector Foreign Policy Proves It Is Relevant and Has No Alternatives." Astana, January 2023.

10. Bjola, C., and Holmes, M. (eds.). Digital Diplomacy: Theory and Practice. London: Routledge, 2015.
11. Nye, J. S. Soft Power: The Means to Success in World Politics. New York: Public Affairs, 2004.
12. Kello, L. The Virtual Weapon and International Order. New Haven: Yale University Press, 2017.
13. International Telecommunication Union. Global Cybersecurity Index, various editions.
14. Diplomatic Informer. "The Rise of Kazakh Diplomacy: Building Bridges Out of Barriers. Five Recommendations."
15. Insight Turkey. "Kazakhstan's Strategic Priorities in the Turkic World."

EXTERNAL SECURITY THREATS TO GEORGIA: GEOPOLITICAL, TERRITORIAL, AND ECONOMIC DIMENSIONS

Giorgi Tsatsiashvili
Master's Degree, Georgian Defence Forces, Georgia

Abstract

National security is one of the fundamental prerequisites for the stable and peaceful development of a modern state. For Georgia, the protection of sovereignty, territorial integrity, economic resilience, and the security of its population represents a major national priority. Georgia's geographical position at the intersection of Europe and Asia has historically made the country strategically important to regional and global powers while simultaneously exposing it to a range of external security challenges.

This article examines the principal external security threats facing Georgia, with particular emphasis on geopolitical pressure, territorial occupation, the military and political role of the Russian Federation, economic vulnerabilities, and the transformation of contemporary security threats. The article argues that Georgia's external security environment is multidimensional and cannot be adequately understood through a traditional military-security framework alone. The continued occupation of Abkhazia and the Tskhinvali region/South Ossetia, the presence of Russian military forces, restrictions on freedom of movement, illegal detentions, borderization, economic pressure, and wider geopolitical competition have significant implications for Georgia's sovereignty and long-term development.

At the same time, economic diversification, international cooperation, diplomatic engagement, peaceful conflict-resolution mechanisms, and institutional resilience provide important instruments for reducing Georgia's vulnerabilities. The article concludes that Georgia's long-term security requires a comprehensive strategy combining territorial resilience, economic diversification, democratic institutional development, international cooperation, technological security, and peaceful conflict resolution. Such an approach is essential for protecting Georgia's sovereignty and supporting its sustainable development in an increasingly complex regional and international security environment.

Keywords: Georgia; national security; external security threats; Russian Federation; territorial integrity; occupied territories; Abkhazia; Tskhinvali region; South Ossetia; economic security; geopolitics; conflict resolution.

1. Introduction

National security is one of the most important challenges facing any modern state. The protection of sovereignty, territorial integrity, economic potential, and the safety of citizens constitutes a fundamental responsibility of the state. Without an effective national security system, sustainable economic development, political stability, and the long-term welfare of society become considerably more difficult to achieve.

For Georgia, the importance of national security is particularly pronounced. The country is located in a strategically important part of the South Caucasus, at the intersection of Europe and Asia and along major transportation and energy routes. Georgia's geographical position provides

significant economic and geopolitical opportunities, but it has also historically exposed the country to competition among larger powers and to various forms of external pressure.

The restoration of Georgia's independence in 1991 created new opportunities for sovereign state development, but the post-Soviet period was also accompanied by significant political instability, armed conflicts, economic difficulties, and unresolved territorial disputes. The conflicts in Abkhazia and the Tskhinvali region/South Ossetia remain among the most important challenges to Georgia's sovereignty and territorial integrity.

The August 2008 war between Russia and Georgia represented a critical turning point in Georgia's modern security environment. The consequences of the war remain relevant to the country's security, sovereignty, territorial integrity, and humanitarian situation. In August 2026, the European Union again stated that the consequences of the 2008 war continue to undermine Georgia's security, sovereignty, and territorial integrity, and reaffirmed its opposition to Russia's military presence in Abkhazia and South Ossetia.

European External Action Service However, Georgia's external security challenges are not limited to conventional military threats. Contemporary security increasingly includes economic, technological, cyber, informational, energy, and political dimensions. External pressure can therefore be exercised without direct military confrontation.

The purpose of this article is to examine the principal external security threats facing Georgia and to analyze their impact on the country's territorial integrity, political stability, economic resilience, and long-term development.

The central argument of the article is that Georgia's national security cannot be ensured exclusively through military capabilities. A comprehensive security strategy must combine territorial defense, diplomatic engagement, economic resilience, international cooperation, institutional capacity, and peaceful conflict-resolution mechanisms.

2. Methodological Approach

The article applies a qualitative analytical approach based on the analysis of official government documents, international institutional statements, security-related reports, and academic literature concerning Georgia's geopolitical and security environment.

The analysis focuses on three interconnected dimensions of external security: territorial and military security, economic security, and broader geopolitical resilience. Particular attention is paid to the unresolved conflicts in Abkhazia and the Tskhinvali region/South Ossetia and to the role of the Russian Federation in Georgia's security environment.

The study also examines the relationship between external security and economic resilience. This relationship is important because political and military pressure may have direct economic consequences, while economic vulnerabilities may in turn increase a state's exposure to external influence.

The methodological approach is therefore based on the assumption that national security should be understood as a multidimensional phenomenon rather than exclusively as the state's ability to respond to conventional military threats.

3. Georgia's Geopolitical Environment

Georgia's geopolitical position is one of the principal factors shaping its national security environment. The country is situated between Europe and Asia and provides access to the Black Sea and important transportation and energy corridors.

The strategic importance of Georgia has increased as international trade, energy security, regional connectivity, and transportation networks have become increasingly important to global and regional actors. Georgia's geographical position provides opportunities for economic

development and international cooperation, but it simultaneously increases the country's exposure to geopolitical competition.

The South Caucasus has historically been characterized by competition among larger powers. The dissolution of the Soviet Union transformed the regional political environment but did not eliminate geopolitical rivalry. Instead, independent states were required to develop foreign and security policies within a complex environment characterized by asymmetric power relations.

For Georgia, this environment has been particularly challenging. The country shares borders with Russia, Azerbaijan, Armenia, and Turkey and has access to the Black Sea. Its geographical position makes it an important link between different economic and political regions.

At the same time, Georgia's proximity to Russia and the existence of unresolved conflicts have created structural security vulnerabilities. The country's security environment is therefore influenced by both its strategic opportunities and its geopolitical risks.

The 2008 war further demonstrated the significance of Georgia's geopolitical position. Following the conflict, the European Union mediated the Six-Point Agreement of 12 August 2008 and established the European Union Monitoring Mission in Georgia. The mission began its monitoring activities on 1 October 2008 and continues to operate along the administrative boundary lines adjacent to Abkhazia and South Ossetia.
European External Action Service+1

Georgia's geographical position should therefore be viewed as both a strategic asset and a potential source of vulnerability.

4. Territorial Integrity as a Central Security Challenge

The unresolved status of Abkhazia and the Tskhinvali region/South Ossetia represents one of the most serious challenges to Georgia's national security.

Georgia's territorial integrity is recognized within its internationally recognized borders, while the two regions remain outside the control of the Georgian central government. Russia maintains a military presence in both territories and has established extensive political, security, and economic relations with the de facto authorities.

The continuation of the Russian military presence creates a persistent security challenge. It limits Georgia's ability to exercise effective control over its internationally recognized territory and creates the possibility of further escalation during periods of regional instability. The security problem also has a significant humanitarian dimension. Communities living close to the administrative boundary lines have experienced restrictions on movement, illegal detentions, the installation of fences and other physical barriers, and difficulties associated with access to education, property, and family relations. The European Union has repeatedly expressed concern over military build-up, borderization, restrictions on freedom of movement, and arbitrary detention in the conflict-affected areas. European External Action Service The European Union's 2026 statement on the anniversary of the 2008 war reaffirmed that Russia's ongoing military presence and activities in Abkhazia and South Ossetia without the consent of the Georgian government are contrary to international law and Russia's obligations under the 12 August 2008 Six-Point Agreement. European External Action Service

The territorial issue is therefore not simply a question of borders. It is simultaneously a military, political, humanitarian, diplomatic, and socioeconomic challenge.

5. The Russian Federation and Georgia's Security Environment

The Russian Federation represents the most significant external military and geopolitical challenge to Georgia's national security.

The history of Russian-Georgian relations demonstrates that political disputes between the two states have frequently had broader security and economic consequences. Following the

dissolution of the Soviet Union, Russia maintained significant political, military, and economic influence in the South Caucasus.

The conflicts in Abkhazia and South Ossetia became central elements of Georgia's post-Soviet security environment. The 2008 war fundamentally changed the nature of the conflict and strengthened Russia's military and political presence in the two regions.

Russia subsequently recognized Abkhazia and South Ossetia as independent entities, a decision rejected by Georgia and by the overwhelming majority of the international community. The European Union continues to reject Russia's recognition of the two occupied regions and supports Georgia's territorial integrity within its internationally recognized borders. European External Action Service+1

The Russian military presence is accompanied by efforts to deepen political, regulatory, security, and economic integration with the occupied territories. The European Union has expressed concern over increasing attempts to integrate Abkhazia and South Ossetia into Russia's regulatory and security space. European External Action Service

From the perspective of Georgia's national security, these processes increase the difficulty of achieving a peaceful resolution of the conflicts and create long-term risks to the restoration of territorial integrity.

Nevertheless, Georgia's declared policy remains focused on peaceful conflict resolution. Diplomatic mechanisms, international mediation, confidence-building measures, and engagement with international organizations remain central instruments of this policy.

6. Human Security and the Consequences of Occupation

A comprehensive analysis of Georgia's security environment should not be limited to state-level military threats. The concept of human security is also essential.

People living in conflict-affected areas may face security risks that do not take the form of conventional military confrontation. Restrictions on movement, detention, barriers to education, property-related disputes, and difficulties in maintaining family and social relations can have long-term effects on communities.

The continuation of such problems contributes to the fragmentation of society and makes reconciliation more difficult.

The European Union Monitoring Mission plays an important role in reducing these risks. The mission is an unarmed civilian mission established after the 2008 war and conducts continuous patrols along the administrative boundary lines. Its objective is to contribute to stability and facilitate a safer and more normal life for conflict-affected populations. European External Action Service

International mechanisms such as the Geneva International Discussions and the Incident Prevention and Response Mechanisms also provide platforms for communication and risk reduction.

These mechanisms cannot by themselves resolve the underlying political conflict, but they can reduce the likelihood of escalation and contribute to confidence-building.

Therefore, Georgia's security policy should treat territorial integrity and human security as closely interconnected objectives.

7. Economic Security as an External Security Dimension

Economic resilience has become an increasingly important component of national security. A state's ability to withstand external pressure depends partly on the strength and diversification of its economy. Excessive dependence on a single external market, energy supplier, transportation route, or investment source may create vulnerabilities that can be exploited during periods of political tension.

Georgia's post-Soviet experience illustrates the relationship between political conflict and economic security.

Following the restoration of independence in 1991, Georgia experienced severe economic difficulties and political instability. Relations with Russia subsequently became an important factor affecting Georgia's economic environment.

The events of 2006 demonstrated the potential connection between political disputes and economic pressure. Russia imposed significant restrictions on Georgian imports and interrupted supplies of natural gas and electricity. Restrictions on trade and the deportation of Georgian citizens further demonstrated the extent to which political tensions could produce economic and humanitarian consequences.

The experience of the 2006 crisis reinforced the importance of economic diversification for Georgia's national resilience.

The 2008 war created additional economic and investment-related challenges. Military conflict and territorial occupation affected infrastructure, investor confidence, economic expectations, and the country's broader development environment.

Georgia has subsequently pursued policies aimed at diversifying trade, energy sources, transportation routes, and investment partnerships.

Economic diversification should therefore be understood as more than a development objective. It is also an element of national security

.

8. Energy Security and Economic Diversification

Energy security is particularly important for countries that operate in geopolitically sensitive regions.

Dependence on a single external energy supplier may create political and economic vulnerabilities. Diversification of energy sources and supply routes can therefore increase a state's ability to resist external pressure.

Georgia has sought to strengthen its role in regional energy and transportation networks and to develop cooperation with multiple external partners.

The development of transportation corridors connecting Europe and Asia also provides Georgia with an opportunity to transform its geographical position into an economic and strategic advantage.

The development of ports, railways, highways, logistics infrastructure, and energy transit routes can contribute to economic resilience while strengthening Georgia's strategic importance to international partners.

However, economic connectivity alone cannot guarantee security. Connectivity must be supported by strong institutions, infrastructure protection, regulatory capacity, and diversified economic partnerships.

Georgia should therefore continue to pursue an economic-security strategy based on diversification rather than excessive dependence on any single external actor.

9. Transformation of Contemporary Security

The nature of external security threats has changed significantly in the twenty-first century.

Traditional military threats remain fundamental, but modern states increasingly face hybrid forms of pressure that combine political, economic, technological, informational, and military instruments.

Georgia is particularly exposed to these forms of pressure because of its geopolitical environment.

Cybersecurity has become increasingly important because critical infrastructure, government institutions, financial systems, transportation networks, and communication systems depend on digital technologies.

Information security is also an important component of national resilience. Disinformation and foreign information influence can affect public perceptions, political processes, and social cohesion.

The European Union has emphasized the importance of strengthening Georgia's resilience against hybrid and cyber threats, including cooperation in strategic communication and countering disinformation. European External Action Service

The changing character of security requires Georgia to expand its understanding of national defense.

National resilience should include:

- territorial defense;
- economic diversification;
- energy security;
- cybersecurity;
- protection of critical infrastructure;
- institutional resilience;
- information security;
- democratic governance;
- social cohesion; and
- international cooperation.

Such an integrated approach can help reduce vulnerabilities that cannot be addressed through conventional military capabilities alone.

10. International Cooperation as an Instrument of National Security

International cooperation represents an essential component of Georgia's national security policy.

As a relatively small state operating in a region characterized by significant power asymmetries, Georgia cannot rely exclusively on its own capabilities. International partnerships provide diplomatic, economic, institutional, and security-related advantages.

The European Union has played a particularly important role in Georgia's security environment. Following the 2008 war, the EU mediated the Six-Point Agreement and established the European Union Monitoring Mission. European External Action Service

The EU continues to support Georgia's sovereignty and territorial integrity and participates in conflict-resolution and confidence-building mechanisms.

International cooperation also contributes to economic resilience. Trade agreements, investment partnerships, transportation projects, and energy cooperation can help Georgia diversify its external relations.

Georgia's cooperation with international organizations should therefore be understood as a strategic component of national resilience rather than simply as a foreign-policy objective.

11. Divergent Perspectives within Georgian Society

Georgia's external security environment has also generated different perspectives within Georgian society concerning foreign policy, economic development, and relations with Russia.

A segment of Georgian society argues that closer relations with Russia could reduce political tensions and provide additional economic opportunities. Supporters of this position may emphasize Georgia's limited military and economic capabilities and the importance of avoiding direct confrontation with a significantly stronger neighboring state.

Another perspective emphasizes the importance of sovereignty, territorial integrity, economic diversification, and deeper cooperation with European and other international institutions.

These different approaches reflect competing assessments of Georgia's strategic interests. The debate demonstrates that national security is not only a question of military capabilities or foreign policy. It is also connected to domestic political consensus, public perceptions, economic expectations, and social resilience.

For Georgia, the challenge is to develop a security policy capable of protecting national interests while minimizing the risks associated with geopolitical confrontation.

A sustainable national security strategy should therefore be based on transparent institutions, public communication, evidence-based decision-making, and broad societal resilience.

12. Georgia's Peaceful Conflict-Resolution Policy

Despite the persistence of the territorial conflicts, Georgia continues to emphasize peaceful conflict resolution.

Military escalation would carry substantial political, humanitarian, and economic risks. For this reason, diplomatic mechanisms and international engagement remain essential.

Georgia participates in the Geneva International Discussions, which provide an international platform for addressing security and humanitarian consequences of the conflicts.

Confidence-building mechanisms are also important. Cooperation in humanitarian issues, freedom of movement, access to education, family contacts, and other practical matters can contribute to the gradual rebuilding of trust.

The objective of conflict-resolution policy should therefore extend beyond the formal restoration of territorial integrity. Sustainable peace requires conditions in which communities can coexist safely and develop economic and social relations.

The long-term restoration of territorial integrity is more likely to be sustainable if accompanied by reconciliation, confidence-building, economic development, and institutional integration.

13. Toward a Comprehensive National Security Strategy

The analysis demonstrates that Georgia requires a comprehensive approach to national security.

First, Georgia should continue to protect its sovereignty and territorial integrity through peaceful and diplomatic means. The conflicts in Abkhazia and the Tskhinvali region/South Ossetia should remain a central priority of national policy.

Second, the country should continue to strengthen economic resilience through diversification. Trade, investment, energy supplies, and transportation partnerships should be diversified in order to reduce exposure to external economic pressure.

Third, Georgia should strengthen institutional resilience. Effective public institutions, the rule of law, accountable governance, and democratic stability are important components of national security.

Fourth, cybersecurity and technological resilience should become increasingly important elements of national defense planning. Critical infrastructure and essential public services require protection against cyber and other non-conventional threats.

Fifth, Georgia should strengthen societal resilience. Social cohesion, public trust, effective communication, and resistance to disinformation can reduce the impact of external influence.

Sixth, international cooperation should remain a central pillar of Georgia's security policy. Diplomatic support, international monitoring, economic partnerships, and cooperation with international organizations can significantly strengthen the country's resilience.

Finally, Georgia's national security strategy should be based on long-term planning rather than short-term responses to individual crises.

14. Discussion

Georgia's experience demonstrates that external security threats are interconnected. Military pressure can affect investment and economic confidence. Territorial occupation can generate humanitarian and social problems. Economic dependence can increase political vulnerability. Political instability can weaken institutional resilience. Information influence can deepen social polarization.

For this reason, external threats should be analyzed as part of a broader security system rather than as isolated problems.

The Russian military presence in Abkhazia and South Ossetia remains the most significant external military challenge to Georgia. However, the country's long-term security will also depend on its ability to strengthen economic, institutional, technological, and societal resilience.

Georgia's geographical position presents both opportunities and risks. The same location that makes Georgia strategically important for international transportation and energy connectivity can also expose it to geopolitical competition.

The most effective response is therefore not isolation but diversification. Diversification of economic relations, energy supplies, transportation routes, investment sources, diplomatic partnerships, and security cooperation can reduce Georgia's vulnerability to pressure from any single external actor.

At the same time, diversification must be accompanied by strong institutions. Economic and geopolitical resilience cannot be sustained without effective governance and social cohesion.

The peaceful approach to conflict resolution is another important element. Georgia's commitment to peaceful mechanisms does not eliminate the security risks associated with occupation, but it reduces the likelihood of renewed large-scale conflict and strengthens international support for Georgia's territorial integrity.

15. Conclusion

Georgia's external security environment is complex, multidimensional, and closely connected to broader geopolitical developments in the South Caucasus.

The most significant external security challenge remains the unresolved territorial conflicts in Abkhazia and the Tskhinvali region/South Ossetia and the continued Russian military presence in these territories. The consequences of the 2008 Russia-Georgia war continue to affect Georgia's security, sovereignty, territorial integrity, and humanitarian environment. The European Union continues to support Georgia's territorial integrity within its internationally recognized borders and opposes Russia's military presence in the occupied regions. European External Action Service+1

At the same time, Georgia's security cannot be understood exclusively through a military perspective. Economic security, energy diversification, cybersecurity, information security, institutional resilience, and international cooperation have become increasingly important components of national security. Georgia's post-Soviet experience demonstrates that external political pressure can have significant economic consequences. The experience of trade restrictions, energy-related pressure, and the economic consequences of armed conflict illustrates the importance of diversification. Economic resilience should therefore be regarded as an essential component of national defense.

Georgia's international partnerships also provide important mechanisms for strengthening resilience. The European Union Monitoring Mission, international diplomatic mechanisms, and

broader international cooperation contribute to stability and maintain international attention on the unresolved conflicts. European External Action Service+1

The long-term security of Georgia requires a comprehensive approach based on five interconnected principles: territorial resilience, economic diversification, institutional strength, international cooperation, and peaceful conflict resolution.

The restoration of territorial integrity and the achievement of sustainable peace cannot be accomplished through a single instrument. They require long-term diplomatic engagement, international support, economic development, societal resilience, and strong national institutions.

Ultimately, Georgia's ability to withstand external threats will depend not only on military capabilities but also on the strength of its economy, institutions, international partnerships, and society.

A sovereign, resilient, democratic, and economically diversified Georgia will be better positioned to manage regional and global challenges while protecting its national interests and pursuing sustainable development.

References

European External Action Service. (2026). *Georgia: Statement by the Spokesperson marking 18 years since the outbreak of the war between Russia and Georgia*. Brussels: European Union.

European External Action Service. (2026). *EU Statement on the Secretary General's thirty-third consolidated report on the conflict in Georgia*. Brussels: European Union.

European External Action Service. (2024). *About the European Union Monitoring Mission in Georgia (EUMM).* Brussels: European Union.

European External Action Service. (2024). *Historical context of EUMM in Georgia.* Brussels: European Union.

Government of Georgia. (2022). *Vision 2030 – Development Strategy of Georgia.* Tbilisi, Georgia.

Government of Georgia. (2024). *Government Programme 2025–2028.* Tbilisi, Georgia.

Koiava, R. (2021). *Transformation of Ethnopolitical Conflicts in Historical Perspective.* Tbilisi: Social Justice Center.

State Security Service of Georgia. (n.d.). *Occupied Territories.* Tbilisi, Georgia.

U.S. Embassy in Georgia. (2023). *2023 Investment Climate Statement – Georgia.* U.S. Department of State.

European Union. (2008). *Six-Point Agreement between Georgia and the Russian Federation, 12 August 2008.*

Pedagogical Sciences

THE IMPACT OF INFORMATIZATION ON THE TRAINING OF YOUNG RESEARCHERS IN THE REPUBLIC OF KAZAKHSTAN

Alexander Vitalievich Bashirov
Head of the Laboratory at the Research Institute of the Karaganda University of Kazpotrebsoyuz, Candidate of Technical Sciences
Viktoria Nikolaevna Golovacheva
Doctor of Pedagogical Sciences, Professor at the Abylkas Saginov Karaganda Technical University
Svetlana Borisovna Glazunova
Doctoral Candidate at the Karaganda University of Kazpotrebsoyuz

Abstract This article describes the results of a survey of students at leading universities in the Republic of Kazakhstan on the impact of computerization on students' research work. The research was conducted as part of the Republic of Kazakhstan grant-funded project AP19676691, “An Innovative Model for Improving the Research Work of Students in the Republic of Kazakhstan.”

The article examines the role of research advisors in shaping the professional competence of young researchers in the era of digital transformation.

Key words : student research, respondent survey, trend dynamics, research advisor

From 2023 to 2025, as part of grant project AP19676691 funded by the Ministry of Education and Science of the Republic of Kazakhstan, an innovative model for improving the research work of students at higher education institutions in the Republic of Kazakhstan was developed and studied

In particular, the research examined the impact of a range of socioeconomic conditions on the activity and effectiveness of students’ research work within higher education institutions. [1]

The research methodology was based on the analysis of survey data from students of various majors and academic years at leading higher education institutions in the Republic of Kazakhstan.[2] A total of 2,497 students participated in the survey. This number of respondents made it possible to confirm the reliability of the results obtained and to justify their generalization as applicable to higher education institutions in the Republic of Kazakhstan.

This article describes an assessment of the impact of students’ computer literacy on the effectiveness of the Student Research Center, depending on the year of study.

Table 1 presents the quantitative distribution of respondents from different years at leading universities in the Republic of Kazakhstan.

Table 1: Numerical Breakdown of Students by Academic Year

Academic Year	Number of Students	Percentage
First	771	30.88%
Second	611	24.47%
Third	590	23.63%
Fourth	525	21.03%
Total	2497	100%

Figure 1 shows trends in students' opinions regarding the impact of computerization on preparing young researchers for research work at universities.

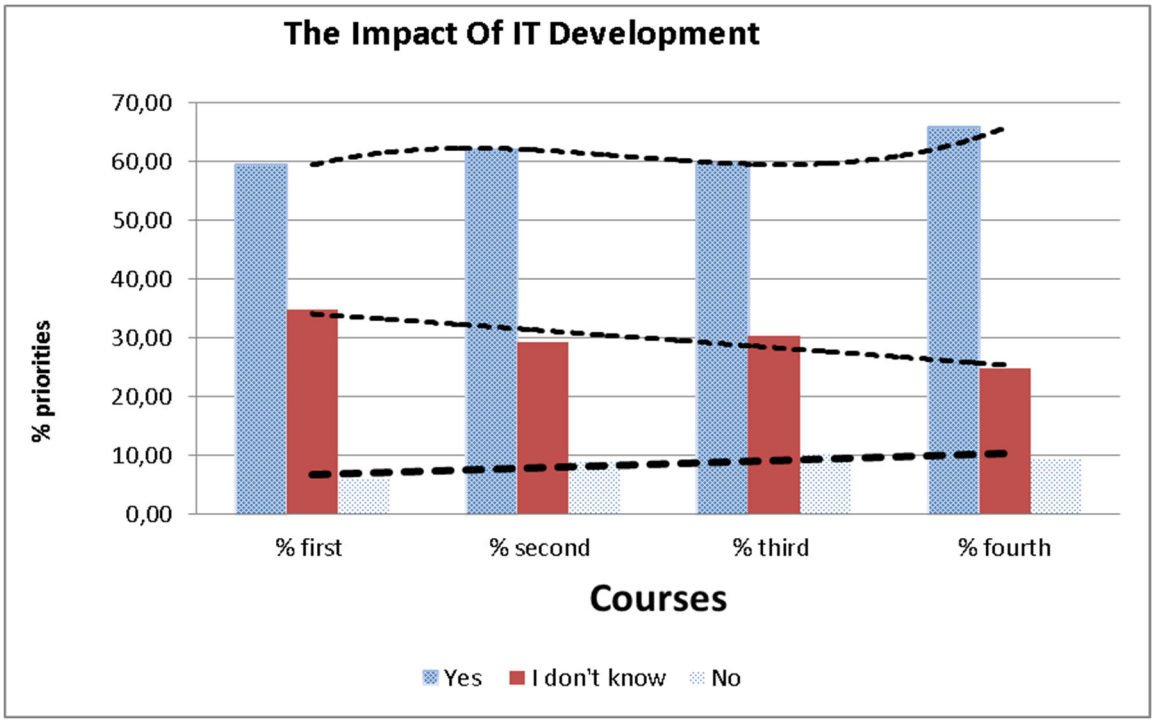


Figure 1
Results of a Student Survey on the Impact of Computerization on the Effectiveness of Undergraduate Research, by Academic Year

Figure 1 shows that the responses of the majority of students (about 60%) confirm that computerization has a significant impact on preparing young researchers for scientific research activities.

The responses of approximately 10% of students deny any influence of informatization, and about 30% express doubts about it.

It is worth noting the consistent trends in the distribution of affirmative "Yes" responses, doubtful "I don't know" responses, and negative "No" responses across academic years.

The trend for "Yes" responses is not continuous; a peak is observed in the second year, a trough in the third year, and then another peak in the fourth year.

The trend for "I don't know" responses is downward, while the trend for "No" responses is upward.

We consider it appropriate to discuss the patterns identified.

In their first year, students are introduced to the fundamentals of academic work and develop skills in searching for and analyzing information.

Many first-year students arrive with varying levels of digital literacy. The initial period is often spent bringing their basic information-handling skills up to par. Students often tend to use the texts they find without fully comprehending them. On the one hand, digitalization facilitates access to necessary resources; on the other hand, it encourages a superficial approach[3].

n their second year, students primarily learn basic methods of data collection and processing. During this period, there is a sharp increase in the use of software tools. Students are often confused about the necessity, appropriateness, and functionality of a particular tool [4,5]

In their third year, students experience information overload. This aligns with the dynamics of trends (Figure 1). During this period, the role of computerization in the development of a young researcher diminishes. It is precisely at this point that interaction with the research advisor becomes extremely important. During this period, students need qualified guidance that includes a detailed research plan and limits on the scope of their research [6].

A fourth-year student's education involves completing a senior thesis. Leveraging the opportunities offered by digitalization enhances the quality and competitiveness of the research conducted [7].

In conclusion, it should be noted that informatization has an undeniable impact on the competence of young researchers during their university education. During their studies, students acquire useful skills in searching for necessary materials, organizing them, processing relevant information, and utilizing the capabilities of computer modeling and artificial intelligence.

The main characteristic is the abundance of available information, as well as the abundance of software tools that have been mastered for processing and organizing information. In this case, the crucial role of mentors and academic advisors is evident.

References

1) Bashirov A.V., Borisov A.A., Garipova A.A., Golovacheva V.N., Glazunova S.N., Smirnov L.S., Tapenova G.S., Hanov T.A. Innovacionnaya model' sovershenstvovaniya nauchno-issledovatel'skoj raboty studentov v Respublike Kazahstan: Monografiya. – Karaganda, 2025. – 300 s. ISBN 978-601-235-734-9.

2) Hanov T.A., Bashirov A.V. Nauchno-issledovatel'skaya rabota studentov v VUZe: prichiny snizheniya aktivnosti//Sovremennye naukoemkie tekhnologii. 2021. № 6-1. S. 209-214.

3) Škrlec, K., & Čančarević, M. (2023). Influence of Information Technologies on the Quality of Study Programs in Higher Education. In M. Van Wyk (Ed.), Technology in Learning. IntechOpen. DOI: 10.5772/intechopen.108647

4)Spivakovska, Y., Vinnyk, T., Perminova, L., & Kotkova, V. (2019). ICT in Professional Education of Future Primary School Teachers: Modeling of Scientific and Research Work. CEUR Workshop Proceedings, Vol. 2393.

5)Xujakulov, Sh. A., & Aminov, I. B. (2020). Formation of Research Competence of Students on the Basis of Information Technologies. JournalNX — A Multidisciplinary Peer Reviewed Journal, ISSN 2581-4230.

6)Oguguo, B. C. E., Okeke, A. O., Dave-Ugwu, P. O., et al. (2020). Assessment of ICT Skills Relevant for Effective Learning Possessed by Undergraduate Students at University of Nigeria. International Journal of Higher Education, 9(4), 206–214. DOI: 10.5430/ijhe.v9n4p206.

7)Anhwere, D. K., Attipoe, E. K., Sayibu, A.-G., Nyaaba, R. A., & Okyere-Darko, A. (2025). Impact of Year-Level Progression on Students' Performance: A Longitudinal Study of LMS Effectiveness at University of Cape Coast (UCC). Creative Education, 16, 428–451. DOI: 10.4236/ce.2025.163027.

THE EVOLUTION OF LINGUACULTURAL COMPETENCE IN MODERN PEDAGOGY

Alima Berzhan
Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
Perizat Yelubayeva
Al-Farabi Kazakh National University, Almaty, Kazakhstan

Annotation

This article presents a comprehensive theoretical review devoted to the study of the genesis, conceptual and categorical apparatus, and key stages in the evolution of linguacultural competence in modern pedagogical science. The relevance of this work stems from the need for a thorough theoretical and methodological understanding of the transformations of foreign language education in the context of globalization processes, the expansion of intercultural dialogue, and the practical implementation of the state trilingualism policy in Republic of Kazakhstan. The main goal of the study is to systematically structure the accumulated scientific and pedagogical knowledge regarding the content, component composition, and promising vectors for the development of this phenomenon.

This article consistently distinguishes between the concepts of "sociocultural", "intercultural", and "linguacultural" competencies, revealing their hierarchical interrelationships, where the linguacultural layer serves as the highest integrating level of language acquisition as a value-semantic code of culture. Based on a comparative analysis, the specific priorities of foreign, domestic, and Kazakhstani research traditions are characterized. The leading trends in the contemporary transformation of this concept are formulated and substantiated: a conceptual transition from factual linguistic and cultural studies to a value-oriented, culture-appropriate approach, as well as the adaptation of linguacultural theory to the conditions of digitalization of the educational environment and online communication. The theoretical significance of the results lies in the systematization of the terminological basis of linguadidactics, which creates the methodological prerequisites for the modernization of curricula, specialized courses, and methodological complexes in language disciplines in schools and universities.

Key words: linguacultural competence, multilingual education, culture-appropriate approach, intercultural communication, linguadidactics, evolution of pedagogical concepts, digital linguacultural studies.

Introduction

The first President of the Republic of Kazakhstan Nursultan Nazarbayev announced the trilingualism project in October 2006 at the 12th session of the Assembly of the People of Kazakhstan. In 2007, in his annual address "New Kazakhstan in a New World", he proposed a phased implementation of this cultural project. He noted that Kazakhstan should be perceived worldwide as a highly educated country whose population speaks three languages: Kazakh as the state language, Russian as the language of interethnic communication, and English as the language of successful integration into the global economy [1, p. 632].

In turn, the current head of state adheres to a similar plan for the development of the linguistic and cultural sphere:

"This is a very complex and important issue. My position is: first, there should be Kazakh and Russian. They are very important for our children. And only then should we teach English", said the President of Kazakhstan [2].

However, large-scale transformation processes in the modern education system demonstrate that the practical implementation of a multilingual model constantly encounters methodological barriers. Simply developing language and speech skills no longer meets the demands of a globalized world. Developing the linguistic and cultural competence of students and future teachers- that is, mastering language as a cultural code, a vehicle of national values, and a reflection of the ethnocultural worldview- is becoming increasingly important.

The problem with this literature review lies in a significant contradiction. On the one hand, there is a high state and societal demand for training specialists with deep linguistic and cultural competence in a trilingual environment, while on the other, pedagogical science still lacks a sufficiently systematized theoretical understanding of the evolution of this concept.

The relevance of this work stems from the need to understand the theoretical foundations of the linguacultural approach to education. In Kazakhstani and global pedagogy, the study of native and foreign languages without reference to national culture often leads to a "cultural blind spot". This seriously hinders intercultural dialogue and harmonious inculturation of the individual.

Level of scientific development. Fundamental ideas about the relationship between language and culture were laid out in the works of W. von Humboldt, Sapir, and Whorf. In Russian linguodidactics and pedagogy, the linguaculture and linguacultural approaches were developed by Vereshchagin, Kostomarov, Vorobyov, Maslova, Ter-Minasova, and Passov. The foreign tradition of intercultural competence is represented by the studies of Byram and Kramsch. In Kazakhstan science, issues of multilingual education and culture-appropriate learning are reflected in the works of Zhetpisbayeva, Akhmetzhanova, Kurmanbayeva, and other authors. At the same time, the genesis and transformation of linguacultural competence under the influence of digital and sociocultural trends require a more systematic retrospective and prospective analysis of existing sources.

The object of analysis is the pedagogical theory and methodology for developing linguacultural competence. *The subject* is the genesis, conceptual and categorical apparatus, substantive evolution, and modern conceptual models of linguacultural competence in pedagogical science.

The purpose of this literature review is to systematize and generalize the existing theoretical approaches to understanding the evolution of linguacultural competence in modern pedagogy.

To achieve this goal, the following *tasks are solved:*

1. To clarify the conceptual and categorical apparatus, identifying the relationships and differences between the concepts of “sociocultural”, “intercultural” and “linguacultural” competence.

2. Conduct a comparative analysis of domestic and foreign scientific schools studying this phenomenon.

3. To identify key vectors of transformation of linguacultural competence in the context of digitalization of the educational space and multicultural environment based on the studied literature.

The scientific novelty of this review lies in its comprehensive systematization of existing theoretical concepts regarding the evolution of linguacultural competence. It also clarifies its component composition, taking into account contemporary sociocultural realities, and reveals a general picture of the transition from traditional linguistic and cultural studies to modern linguacultural models.

The theoretical significance lies in the generalization and structuring of accumulated knowledge about the structure and genesis of linguistic and cultural competence. The results obtained can serve as a basis for the further development of curricula, specialized courses, and textbooks in language disciplines in schools and universities.

Methods and materials

This literature review draws on the research of Russian and international authors on issues of cultural linguistics, linguadidactics, intercultural communication, and multilingual education. The work includes monographs, research articles, dissertations, and textbooks published from the mid-20th century to the present. Particular attention was paid to Kazakhstani sources reflecting the specific implementation of state language policy. In total, more than 30 sources were analyzed, including works by classics and contemporary researchers.

The primary methods employed were a combination of theoretical approaches typical of literary analysis in pedagogy. Primarily, systemic-structural analysis was employed, allowing for the examination of linguistic and cultural competence as a holistic system with its components, connections, and stages of development.

The comparative method helped draw parallels between domestic and foreign academic schools, identifying common trends and national characteristics in the interpretation of the concept. Through retrospective analysis, it was possible to trace the historical evolution of the concept from linguistic and regional studies to modern digital interpretations of cultural linguistics.

In addition, a terminological analysis was employed to clarify and differentiate related concepts: "sociocultural competence," "intercultural competence", and "linguacultural competence." To summarize the results, methods of abstraction, synthesis, and systematization of scientific material were used, making it possible to construct a more or less holistic picture of the development of the phenomenon under study based on existing publications.

All conclusions and propositions in this review are theoretical in nature and are based solely on a critical analysis and comparison of literary sources. Empirical methods were not used in this study, as the primary objective was to systematize and synthesize accumulated scientific knowledge.

Results and discussion

An analysis of the scientific literature shows that the development of linguistic and cultural competence in the context of multilingual education in Kazakhstan has been actively discussed over the past two decades. Many authors emphasize that the trinity of languages has become not only a political project but also an important factor in the competitiveness of the country and its citizens [1; 6].

Altynbekova notes the uniqueness of the linguistic situation in the republic, where different types of language policy are simultaneously implemented: vernacularization, monolingualism, multilingualism, and internationalization. The researcher characterizes the Kazakh model as balanced, as it combines the objectives of preserving the state language with supporting multilingualism and integration into the global space [3, pp. 86-91].

Considerable attention in the literature is devoted to the relationship between language and culture. The works of Bazylova, Mamedova, Dulayeva emphasize that linguacultural competence is a key element of intercultural communication, as its subject matter closely intersects with the study of ethnic mentality. The authors examine in detail the specific features of Kazakhstan as a multi-ethnic state, where more than one hundred nationalities maintain their own identity while simultaneously participating in the formation of a common Eurasian culture [1 ;4].

The idea of Eurasianism as a synthesis of Western and Eastern values has been widely covered in the works of Kazakhstan scholars (Sabitova, Zhuravleva, Sklyarenko, Shetieva, and others). In their opinion, the long-term coexistence of peoples has formed a special Eurasian mentality, which facilitates intercultural interaction due to a common cultural code and system of values [5, pp. 80-85; 6].

Contemporary research notes that, despite the successes of public policy, a number of problems remain. These include the risk of folklorization of the state language, linguistic interference, and the prevalence of economic expediency over the cultural dimension of education [7; 8]. Many authors point out that multilingual education is currently focused primarily on the acquisition of individual languages rather than on the comprehensive development of linguacultural competence as a holistic phenomenon [1; 9].

A wide range of researchers (Kenzhebekov, Maslov, Prokhorov, Vereshchagin, Kostomarov, Vorobyov, Kebekov, Konetskaya, Maltsev, Safonova, Sysoev, Tomakhin) emphasizes the fundamental importance of the linguacultural approach. Their works consistently substantiate the thesis that mastering a foreign language is impossible without a parallel understanding of the world, mentality, and culture of its speakers [9; 10; 11].

In the context of studying the communicative functions of language, Jakobson's model is of fundamental importance. [19], according to which the linguistic act inevitably links the context, message and cultural-semiotic code. We agree with Galskova and. Gez, proving the importance of mastering the skills of interaction in an intercultural environment and emphasizing that otherwise “communicators are, as it were, “programmed” in advance for a conflict of misunderstanding” [10].

An analysis of the literature allows us to identify several key approaches to understanding and developing linguistic and cultural competence:

1. The communicative approach (Bugam, Kramsch and their followers) focuses on the actual use of language in an intercultural context and the formation of an “intercultural speaker”.
2. The cultural approach emphasizes that language is an integral part of culture and social experience (Vereshchagin, Kostomarov, Ter-Minasova, Maslova).
3. The personality-oriented approach, based on the ideas of humanistic psychology, highlights the autonomy of the learner and takes into account his individual characteristics [8; 12; 13; 14].

Kazakhstan researchers are actively developing these ideas in relation to the national context. The works of Zhetpisbayeva, Akhmetzhanova, Kurmanbayeva, Zhumasheva, Demessinova, Beisembayeva, Mekezhanova, Zhetpisbay and others demonstrate the specifics of implementing multilingual education in schools and universities, the problems of training teaching staff and ways of integrating the linguacultural component into the educational process [7; 9; 15; 16; 17].

A separate section of literature is devoted to the conceptual and structural transformation of linguacultural competence and the readiness of students to master it. Theoretical studies show that mastering linguacultures requires not simply the acquisition of vocabulary, but a holistic psychological readiness of the individual for complex speech and intercultural activity (Dyachenko, Kandybovich [23]). Understanding language as a semantic space is based on the semiotic concept of culture by Yu. M. Lotman [20], where language acts as the code of the "semiosphere".

In domestic pedagogy, this synthesis is reflected in the foreign language educational concept of Kunanbaeva [21], which proposes the modernization of foreign language education based on a linguacultural communicative approach that ensures the formation of the value-semantic core of the student’s personality.

Thus, the need arises for a precise delineation of the conceptual and categorical apparatus of the area under study, in which basic categories are often used synonymously. Sociocultural competence in the works of Safonova V. V. and Sysoeva P.V. is interpreted primarily as knowledge of the sociocultural context, rules of conduct, and national etiquette. Intercultural competence, according to Galskova's research., Gez [10] and Byram shifts the focus to the individual’s readiness for practical dialogue and overcoming communication failures.

In turn, linguacultural competence represents an integrating level, which presupposes mastery of language as a cultural code through the comprehension of concepts and precedent texts [11; 21]. In understanding the communicative-semiotic functions of language, Jakobson's model acquires key significance. R. [19], linking the linguistic act with the cultural code, as well as the provisions of Dyachenko and Kandybovich [23] on psychological readiness for speech activity within the framework of the “semiosphere” of Yu. M. Lotman [20].

Below is a table to visually confirm the results of the study and compare linguistic and cultural competence and its evolution. A comparison of leading scientific schools reveals the presence of several research vectors, differing in their goals and methodological emphases.

Table 1 – Comparative table of key areas linguacultural competence in the educational process

Scientific school / Tradition	Main representatives	Key research focus
Foreign (Western) school	M. Byram , K. Kramsch , C. Rogers [22]	Formation of an “intercultural speaker”, educational autonomy, facilitation of communication in an international environment.
Domestic (post-Soviet) school	E. M. Vereshchagin, V. G. Kostomarov, V. V. Vorobyov, V. A. Maslova, S. G. Ter-Minasova, G. D. Tomakhin [9; 10; 11]	Study of the value-semantic content of language, national concepts and ethnocultural mentality.
Kazakhstan school	B. A. Zhetpisbaeva [7; 9], O. B. Altynbekova [3], Z. K. Sabitova [5], S. S. Kunanbaeva [21], Z. K. Akhmetzhanova, Sh. K. Kurmanbaeva [16; 17]	Synthesis of global approaches in the context of trilingualism, preservation of cultural code and overcoming language interference.
Note: compiled by the author.		

Structuring the approaches considered allows us to identify two key vectors that determine the evolution of linguacultural competence in modern pedagogical science.

The first vector is culturally appropriate, which is expressed in a shift in focus from classical linguistics and regional studies with its factual memorization to the formation of the value-semantic core of the individual. In Kazakhstani pedagogical thought, this tendency is most fully reflected in the works of Kunanbayeva [21] and Zhetpisbayeva [7; 16], where language acquisition occurs through the prism of a deep understanding of national and universal values.

The second vector of the evolution of linguacultural competence is driven by the widespread digitalization of the educational space and the emergence of digital linguacultural studies. In a virtual environment, the methods of representing cultural codes are changing significantly.

So, traditional Linguacultural units are supplemented by network codes, media texts, and visual symbols. The boundaries of access to an authentic communicative environment are expanding. and puts forward new demands on pedagogy: the need to develop students' skills for critically decoding and interpreting digital cultural meanings [18; 21]. Thus, the conceptual transformation of this phenomenon reflects the transition from a rigid paradigm to flexible, modeled competencies that meet the challenges of a multicultural and digital society.

Conclusion

Summarizing the theoretical approaches discussed, we can identify the main stages and substantive guidelines that determine the development of linguacultural competence in modern pedagogical science. A review of the literature revealed that the conceptual and categorical apparatus in this field has come a long way from a rather utilitarian view of language as a tool to its understanding as a profound value-semantic code of culture. Moreover, the sociocultural, intercultural, and linguacultural components do not overlap, but rather represent successive levels of mastering linguistic reality, where the linguacultural level acts as the highest, integrating form.

A comparative analysis of various research traditions reveals significant differences between academic schools. Foreign authors most often emphasize the development of intercultural speakers and the promotion of personal autonomy in communication. Domestic linguodidactics, in turn, traditionally relies on conceptology, the study of national mentality, and the ethnocultural worldview. Kazakhstan pedagogical science is attempting to find its own path by synthesizing these approaches and adapting them to the realities of trilingual education and a multiethnic society.

At the current stage, the evolution of this concept is characterized by two important shifts. First, there is a definitive shift from simple informative linguistic and cultural studies to culture-specific education, where the primary goal is the development of a linguacultural personality. Second, under the influence of digitalization, traditional linguistic cultures and texts are increasingly being supplemented by virtual codes and new media formats. This poses a new challenge for pedagogy: teaching students to critically understand digital cultural meanings.

Overall, the theoretical foundation accumulated in the literature provides a sound basis for the practical application of the linguacultural approach in real-world educational settings. The review materials may be useful in developing curricula, specialized courses, and teaching aids in language disciplines in both schools and universities, as well as in training future teachers.

Further research in this area should be directed towards the search for and experimental testing of specific methods for developing linguistic and cultural competence in students in a multilingual environment.

List literature

1. Bazylova B., Mamedova, Dulayeva E., Dulayeva N. The Formation of Linguacultural Competence of Students in the Process of Teaching. // Talent Development & Excellence Vol.12, No.1, 2020, 632-645.

2. [Electronic resource]. Tokayev on trilingualism: first Kazakh and Russian, then English. // 2019. URL : https :// tengrinews . kz / kazakhstan news / tokaev - trexieiazycii - snacala - kazaxskii - russkii - angliiskii -365910/ (date of access: 23.07.2026).

3. Altynbekova O. Language policy of Kazakhstan: linguistic proficiency of the largest ethnic groups. // Materials of the international scientific and practical conference “Actual problems of modern philology: theoretical problems and applied aspects” (IV Bagizbaev readings) – Almaty: Kazakh University, 2012, P. 86–91.

4. Dulayeva E., Mamedova F., Bazylova B. Addressing as a Speech Act: Similarities and Differences in Turkish and Arabic. // Messenger Karaganda university Series " Philology ". No. 2(110)/2023. 48-56.

5. Sabitova Z. and oth . Dictionary of Eurasian Linguoculture. Language – Community – Time. Int.conf . "Akhanov readings" under the aegis of MAPRYAL (materials of reports and messages) – Almaty: Kazakh University, 2010, vol. 2, P.80–85.

6. Nazarbayev NA The decree of the President of the Republic of Kazakhstan of June 29, 2011 No. 110 “About the State program of functioning and development of languages for 2011 – 2020”.

7. Zhetpisbayeva BA To the issue of theoretical and methodological conceptualization of polylingual education. Current problems of philology and methodology of foreign language teaching. – 2014. – No. 8. 127 – 135.

8. Adilbayeva VS, Tleuzhanova GK From the experience of development of European and polylingual education. Lingua mobilis . – 2012. – No. 1 (34). – p. 67.

9. Zhetpisbaeva B.A. Theoretical and methodological foundations of multilingual education: Abstract of Doctor of Pedagogical Sciences. — Karaganda, 2009.

10. Galskova N.D., Gez N. I. Theory of teaching foreign languages. // Lingvodidactics and methodology: a teaching aid for students of lingu. universities and faculties of foreign languages of higher pedagogical educational institutions. – 220 p.

11. Vezhbitskaya A. Language, Culture. Cognition, - Moscow: International Relations, 1996. - 240 p.

12. Tazhibayeva S.Zh. , IbrashevaA.Kh . Polylingual education in RK: problems and prospects. – Taraz: Taraz State Pedagogical University.

13. Eshimbetova ZB, Demeuova AM Kazakhstan's "trilingualism" - a unique formula of language policy. Innovation in science: coll. of articles on a material. XI international scientific – practical conference. Part II. – Novosibirsk: SibAK, 2012.

14. Mazhitayeva Sh., Smagulova G., Tuleuova B. Polylingual education as one of the priority directions of development of the system of education in the Republic of Kazakhstan. European Researcher. Series A – 2012. – No. 11-1. – P. 1864 – 1867.

15. Nurgaliyev SA, Kumargaliyeva of ZO Polylinguism in cross-cultural space of Kazakhstan. Open Innovation: collection of articles of the International scientific and practical conference. In 2 parts, Part. 2. – Penza: International Center for Science Education "Science and Education". – 2017. – P. 172-175 ISBN 978-5-907012-24-0 P. 2. ISBN 978-5-907012-22-6

16. Zhetpisbayeva BA, Syrymbetova LS, Kubeeva AE To the issue of training of teachers for multilingual education in Kazakhstan. Current problems of philology and methodology of foreign language teaching. – 2017. – Volume. 11. – P. 168-172.

17. Aitbayeva BM, KadinaZh . Z., Sateyeva BS, Akzhunusova NB, Gornaya A. Role and place of the state language in polylingual formation of the Republic of Kazakhstan. Current problems of humanitarian and natural sciences. – 2015. – No. 12-1. 121-123.

18. IsmagulovaB.Kh., Bazarbayeva A.S., SametovaPh.T. Tendencies of language situation development in Kazakhstan during years of independence. Electronic resource: http://refio.ru/doc_7533a19601838c3567e9c2dbff69a898.html.

19. Jacobson, R., "Linguistics and Poetics", in T. Sebeok , ed., Style in Language, Cambridge, MA: MIT Press, 1960, pp. 350-377.

20. Lotman M. A Semiotic theory of Culture. Indiana University Press, 2000.

21. Kunanbayeva SS The Modernization of Foreign Language Education: The LinguoculturalCommunicative Approach. Hertfordshire Press. – London, UK, 2013. – P. 48-50.

22. Rogers, Carl, Lyon, Harold C., & Tausch, Reinhard. On Becoming an Effective Teacher – Person-centered Teaching, Psychology, Philosophy, and Dialogues with Carl R. Rogers and Harold Lyon. London: Routledge, (2013). – P. 57-60.

23. Dyachenko, M. I. & Kandybovich , L. A. (1976). Psychological problems of readiness for activities. Minsk: Publishing House of the Belarusian University, - P.6.

AI AND CRITICAL THINKING: CAN TECHNOLOGY HELP STUDENTS THINK BETTER?

Bilibayeva Assem
Master's student, International Taraz University named after Sherkhan Murtaza, Master's program in Foreign Languages (Foreign Language Teacher Education).

Abstract
The rapid integration of artificial intelligence (AI) tools, particularly generative large language models such as ChatGPT, into higher education has intensified debate over their effect on students' critical thinking. This article presents a narrative literature review that synthesizes recent empirical and theoretical work (2020–2025) on the relationship between AI tool use and critical thinking development, organized through Facione's (1990) model of critical thinking skills and Anderson and Krathwohl's (2001) revision of Bloom's taxonomy. The review indicates that AI's influence on critical thinking is neither uniformly positive nor uniformly negative but is conditional on how the technology is used. When AI functions as a scaffold that structures inquiry, prompts reflection, and requires learners to verify and justify outputs, it is associated with gains in argumentation, evaluation, and metacognitive awareness. When AI functions as a substitute for independent reasoning, its use is associated with cognitive offloading, reduced neural engagement, and a decline in measured critical thinking scores. The article concludes that AI can help students think better only under deliberate pedagogical conditions and proposes directions for instructional design, assessment, and future research.
Keywords: artificial intelligence; critical thinking; cognitive offloading; higher education; generative AI; Bloom's taxonomy; AI literacy

1. Introduction

Artificial intelligence (AI) has moved from the periphery of educational technology to its center. Generative AI systems such as ChatGPT are now routinely used by students to summarize readings, draft essays, solve problems, and answer study questions (Gonsalves, 2024). This shift has renewed a long-standing pedagogical concern: whether tools designed to reduce cognitive effort also reduce the very thinking they are meant to support. The stakes of this question are high, because critical thinking is widely regarded as one of the central outcomes of higher education and one of the competencies most valued by employers and civic life alike.

Critical thinking is commonly defined, following the Delphi Report of the American Philosophical Association, as "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference" (Facione, 1990, p. 2). Facione's (1990) framework further identifies explanation and self-regulation as core skills, situating critical thinking not merely as a set of logical operations but as a disposition toward reflective, evidence-based reasoning. A complementary lens is offered by Anderson and Krathwohl's (2001) revision of Bloom's taxonomy, which orders cognitive processes from lower-order skills (remembering, understanding, applying) to higher-order skills (analyzing, evaluating, creating). Together, these frameworks provide the conceptual vocabulary used throughout this article to classify how AI tools intersect with different dimensions of student thinking.

The literature on AI and critical thinking has developed along two largely separate lines. One body of work frames AI as a cognitive partner: a source of personalized feedback, alternative

perspectives, and scaffolded practice that can strengthen argumentation and reflection (Gonsalves, 2024). A second, more cautionary body of work frames AI as a source of cognitive offloading, in which the convenience of instant answers displaces the effortful, iterative reasoning that critical thinking requires (Gerlich, 2025). These two narratives are not necessarily contradictory; rather, they may describe different conditions of use. The present article asks: under what conditions does AI use help students think better, and under what conditions does it appear to erode independent reasoning? To address this question, the article synthesizes recent literature using the IMRAD structure, situating the findings within Facione's (1990) and Anderson and Krathwohl's (2001) frameworks and concluding with pedagogical and research implications.

2. Materials and Methods

This article adopts a narrative, integrative literature review design rather than a primary empirical study, which is consistent with the exploratory and synthetic aim of examining a rapidly evolving and heterogeneous evidence base. Narrative review is appropriate when the goal is to integrate findings from studies that differ substantially in methodology, population, and outcome measures — as is the case in current AI-and-education research — and to relate them to established theoretical frameworks.

Sources were identified through targeted searches combining the terms “artificial intelligence,” “generative AI,” or “ChatGPT” with “critical thinking,” “higher-order thinking,” or “cognitive offloading,” restricted to peer-reviewed journal articles, preprints from recognized research institutions, and systematic reviews published between 2020 and 2025. Priority was given to studies involving student or higher-education populations, to mixed-method or quantitative studies with clearly reported measures of critical thinking, and to systematic reviews that themselves synthesized multiple primary studies. Sources that could not be independently verified for authorship or publication details were excluded from citation to preserve the accuracy of the reference list.

The selected literature was organized using two complementary analytical lenses. First, Facione's (1990) six-skill model (interpretation, analysis, evaluation, inference, explanation, self-regulation) was used to identify which specific components of critical thinking each study addressed. Second, Anderson and Krathwohl's (2001) revised Bloom's taxonomy was used to distinguish effects on lower-order thinking skills (remembering, understanding, applying) from effects on higher-order thinking skills (analyzing, evaluating, creating). Findings were then grouped thematically into evidence of facilitative effects, evidence of erosive effects, and evidence of conditional or moderating factors, which form the three subsections of the Results below.

As a narrative rather than systematic review, this study does not claim exhaustive coverage of the literature, nor does it apply formal risk-of-bias scoring to individual studies. This limitation is addressed further in the Discussion.

3. Results

3.1 Evidence that AI use can strengthen critical thinking

Several studies report that AI tools, when embedded in structured learning tasks, are associated with measurable gains in critical thinking. In a longitudinal study of MSc marketing students interacting with generative AI tools over four weeks, Gonsalves (2024) found that AI use engaged cognitive, affective, and metacognitive dimensions of thinking, with students showing gains in melioration, ethical reasoning, collaboration, and reflective thinking when AI-generated content was treated as a starting point for further inquiry rather than a final answer. Gonsalves (2024) further proposed a revised, AI-specific extension of Bloom's taxonomy to capture cognitive processes — such as verifying and re-purposing AI output — that the original taxonomy does not describe.

Systematic reviews of the ChatGPT-and-critical-thinking literature reach broadly convergent conclusions. A systematic literature review of studies on ChatGPT and critical thinking

in EFL (English as a foreign language) writing instruction found that ChatGPT can support argumentative reasoning, encourage learners to weigh multiple perspectives, and provide feedback that facilitates revision, while cautioning that these benefits depend on appropriate pedagogical guidance from instructors (Salsabilla et al., 2026). This conditionality — benefit contingent on structure and guidance — recurs throughout the literature and is examined further in Section 3.3.

At the level of specific cognitive skills, the evidence is most consistent for the lower tiers of Bloom's taxonomy. Gonsalves (2024) reported that AI text generation tools improved performance across Bloom's levels overall, but with greater and more consistent gains for remembering and understanding than for evaluating and creating. This pattern suggests that AI is most reliably beneficial for tasks such as information retrieval, summarization, and comprehension checking, and less reliably beneficial — and potentially counter-productive — for tasks that require independent judgment and original synthesis.

3.2 Evidence that AI use can weaken critical thinking

A second body of evidence documents a negative association between frequent AI use and critical thinking performance, centered on the concept of cognitive offloading — the delegation of cognitive tasks to an external tool in place of internal, effortful reasoning. In the most extensive quantitative study identified in this review, Gerlich (2025) surveyed and interviewed 666 participants across age groups and educational backgrounds, combining standardized critical thinking assessment with thematic analysis of qualitative interviews. The study found a statistically significant negative correlation between frequency of AI tool use and critical thinking performance, with cognitive offloading identified as a mediating mechanism. Younger participants showed both higher AI dependence and lower critical thinking scores than older participants, while higher educational attainment was associated with stronger critical thinking regardless of AI use, suggesting that prior academic training may buffer against offloading effects (Gerlich, 2025).

Complementary neurophysiological evidence comes from a controlled study by Kosmyna et al. (2025) at the MIT Media Lab, in which 54 participants wrote essays under three conditions — using ChatGPT, using a conventional search engine, or writing unaided — while their brain activity was recorded via electroencephalography (EEG) across 32 regions. Participants who used ChatGPT showed the weakest and least distributed neural connectivity of the three groups and, in the researchers' words, “consistently underperformed at neural, linguistic, and behavioral levels” relative to the other groups (Kosmyna et al., 2025). Notably, when ChatGPT users were subsequently asked to write unaided, their brain engagement remained comparatively low, suggesting that reduced cognitive engagement did not immediately reverse once external support was withdrawn. Participants who had relied on ChatGPT across multiple sessions also increasingly resorted to copying AI-generated text with minimal revision, and struggled more than other groups to accurately quote or recall content from essays they had just produced.

Survey research with student populations corroborates these experimental findings at the level of self-perception. In a survey-based study of AI's role in developing critical thinking, Szmyd and Mitera (2024) found that although most students rated their own critical thinking skills as high or moderate, 83% expressed concern that excessive reliance on AI could weaken their capacity for independent thought and responsible decision-making. Students in this study also affirmed that AI could support, but not replace, the traditional teaching methods they saw as essential for developing autonomous reasoning (Szmyd & Mitera, 2024). This combination of continued confidence and simultaneous concern reflects an awareness among students themselves that convenience and cognitive development may be in tension.

3.3 Conditional and moderating factors

Taken together, the facilitative and erosive findings summarized above are not necessarily inconsistent; several factors appear to moderate whether AI use helps or harms critical thinking

development. The first is cognitive level: evidence from Gonsalves (2024) indicates that AI is more reliably beneficial for lower-order thinking skills (remembering, understanding, applying) than for higher-order skills (analyzing, evaluating, creating), where the risk of premature closure — accepting an AI-generated answer without independent evaluation — appears greatest.

The second moderating factor is the structure and intentionality of the task. Reviews of AI use in EFL writing instruction emphasize that benefits to critical thinking depend on deliberate pedagogical guidance — assignments that require students to question, verify, and critically engage with AI-generated responses rather than accept them at face value (Salsabilla et al., 2026). This aligns with Gerlich's (2025) finding that cognitive offloading, not AI use per se, is the operative mechanism linking AI to weaker critical thinking: offloading occurs when a tool is used to bypass reasoning rather than to support it.

The third moderating factor is the learner's existing academic preparation. Gerlich (2025) found that individuals with higher educational attainment maintained stronger critical thinking performance despite frequent AI use, suggesting that a solid foundation in independent reasoning may act as a protective factor against offloading, whereas learners still developing these foundational skills — including many undergraduate and secondary students — may be more vulnerable to its erosive effects. Age showed a related pattern, with younger participants in Gerlich's (2025) sample exhibiting both higher AI dependence and lower critical thinking scores, a finding with direct relevance for the design of AI use policies at different educational levels.

4. Discussion

The evidence reviewed here does not support a single, unconditional answer to the question of whether AI helps students think better. Instead, it supports a conditional answer: AI use can help students think better when it is structured to require verification, justification, and reflection, and it can weaken critical thinking when it substitutes for these processes. This conclusion is consistent with the theoretical distinction, implicit in both Facione's (1990) and Anderson and Krathwohl's (2001) frameworks, between critical thinking as a product (a correct answer or a well-formed argument) and critical thinking as a process (the interpretation, evaluation, and self-regulation that produce that answer). AI tools can readily supply the product; whether they support the process depends on how they are used.

This distinction has direct implications for the revision of instructional frameworks. Gonsalves's (2024) proposed extension of Bloom's taxonomy — which incorporates AI-specific competencies such as verifying, interrogating, and re-purposing AI-generated content — responds to a genuine gap: the original and revised taxonomies were formulated before AI could generate plausible answers at any level of the cognitive hierarchy, and neither offers explicit guidance on how educators should treat AI-assisted performance at the analyzing, evaluating, and creating levels, where the risk of unearned competence is highest.

For classroom practice, the reviewed evidence points toward several concrete strategies. First, assessment design can shift emphasis from the final product to the reasoning process — for example, requiring students to submit not only an AI-assisted answer but also an evaluation of that answer's accuracy, an identification of its limitations, and a justification for any revisions made (a practice consistent with the scaffolding effects reported by Salsabilla et al., 2026). Second, explicit instruction in AI literacy — including how generative models can produce fluent but inaccurate or biased content — may help students maintain the skepticism and verification habits that Gerlich (2025) associates with resistance to cognitive offloading. Third, because younger and less academically prepared students appear more vulnerable to offloading (Gerlich, 2025), institutions may need to calibrate the degree of AI integration to students' developmental stage and prior training in independent reasoning, rather than applying a single policy across all levels of education.

Several limitations qualify these conclusions. Many of the primary studies synthesized here rely on self-reported measures of critical thinking, cross-sectional designs, or relatively small and geographically concentrated samples — limitations acknowledged by the studies' own authors (e.g., Kosmyna et al., 2025, note their sample was drawn from a small number of Boston-area universities). Cross-sectional designs, in particular, cannot establish whether AI use causes changes in critical thinking or whether individuals with weaker critical thinking dispositions are simply more inclined to rely on AI. Longitudinal and experimental designs, such as the controlled EEG study by Kosmyna et al. (2025), offer stronger causal evidence but remain rare, and none of the studies reviewed here followed students over a full academic program. This review is itself narrative rather than systematic, meaning it does not apply standardized inclusion criteria or bias assessment across the full population of relevant studies, and its conclusions should be read as an informed synthesis rather than a definitive quantitative estimate of AI's effect on critical thinking.

Future research would benefit from longitudinal designs that track critical thinking development over full courses or academic years, experimental studies that manipulate the degree of scaffolding built into AI-assisted tasks, and discipline-specific investigations, since the cognitive demands of critical thinking in, for example, language learning, mathematics, and the natural sciences are not identical. Cross-cultural and cross-institutional replication would also help determine whether the moderating role of prior education observed by Gerlich (2025) generalizes beyond the populations studied to date.

5. Conclusion

The question of whether AI can help students think better does not have a categorical answer; the evidence reviewed here indicates that the outcome depends on the manner and purpose of AI use. When generative AI tools are integrated as scaffolds that prompt reflection, require verification, and support argumentation, they are associated with gains in aspects of critical thinking, particularly at the lower and middle tiers of Bloom's taxonomy. When these tools are used as substitutes for independent reasoning, they are associated with cognitive offloading, reduced neural engagement during writing tasks, and measurable declines in critical thinking performance, with younger and less academically prepared learners appearing most vulnerable. These findings suggest that the central pedagogical task is not whether to permit AI in education, but how to design its use so that it augments, rather than replaces, the interpretation, analysis, evaluation, inference, explanation, and self-regulation that constitute critical thinking. Achieving this balance will require deliberate curricular design, explicit AI literacy instruction, and assessment practices that make students' reasoning — not only their answers — visible and accountable.

References

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report). California Academic Press.

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), Article 6. https://doi.org/10.3390/soc15010006

Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. Journal of Marketing Education. Advance online publication. https://doi.org/10.1177/02734753241305980

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task (arXiv:2506.08872). arXiv. https://doi.org/10.48550/arXiv.2506.08872

Salsabilla, N., Munir, S., Safitri, R., & Indayani, V. S. (2026). The role of ChatGPT in enhancing critical thinking in EFL writing: A systematic literature review. Linguavision: Journal of English Education, Linguistics, and Literature, 1(1), 47–63.

Szmyd, K., & Mitera, E. (2024). The impact of artificial intelligence on the development of critical thinking skills in students. European Research Studies Journal, 27(2), 1022–1039. https://doi.org/10.35808/ersj/3876

Historical Sciences

УДК 903(574)

Тюркский археолого-этнографический комплекс «Кумай»: структура и значение в системе историко-культурного наследия Акмолинской области

Turkic Archaeological and Ethnographic Complex “Kumai”: Structural Features and Cultural Significance in the Akmola Region, Kazakhstan

Жаманова Гульмира
магистр гуманитарных наук, заведующий отделением охраны памятников археологии, градостроительства и архитектуры КГУ «Центр по охране и использованию историко-культурного наследия» управления культуры Акмолинской области, г.Кокшетау, Казахстан

Gulmira Zhamanova,
M.A. in Humanities, Head of the Department for Protection of Archaeological Monuments, Urban Planning, and Architecture,
State Institution “Center for the Protection and Use of Historical and Cultural Heritage” of the Department of Culture of Akmola Region, Kokshetau, Kazakhstan

Аннотация

В статье рассматривается тюркский археолого-этнографический комплекс «Кумай», расположенный в Акмолинской области, как объект историко-культурного наследия. Анализируются его структура и основные элементы, включая курганы, могильники и каменные изваяния (балбалы). Рассматривается связь материальной культуры с мифологическим образом Кумай (Умай, Хумай), отражающим представления о защите и продолжении рода. Подчёркивается значение комплексного подхода к изучению памятника и необходимость его сохранения.

Abstract:

The article examines the Turkic archaeological and ethnographic complex “Kumai” in the Akmola Region of Kazakhstan as an important site of historical and cultural heritage. Its structure and main elements, including burial mounds, grave sites, and stone statues (balbals), are analyzed. The study highlights the connection between material culture and the mythological figure Kumai (Umai, Humai), reflecting concepts of protection and lineage preservation. The importance of an integrated interdisciplinary approach and the need for preservation of the site are emphasized.

Ключевые слова

археология Казахстана, тюркская культура, Кумай, балбалы, курганы, Умай, Хумай, культурное наследие

Keywords:
Kazakhstan archaeology, Turkic culture, Kumai, balbals, burial mounds, Umai, Humai, cultural heritage

Сохранение и изучение историко-культурного наследия является важной составляющей духовного развития общества. Акмолинская область обладает уникальными памятниками, отражающими многовековую историю степной цивилизации. Одним из таких значимых объектов является тюркский археолого-этнографический комплекс «Кумай», играющий важную роль в исследовании общетюркской цивилизации.

Тюркский археолого-этнографический комплекс «Кумай» расположен в Ерейментауском районе Акмолинской области. Он был открыт в результате разведочных археологических исследований, проведённых в 2009 году. Согласно отчёту археологических исследований 2010 года, памятники комплекса занимают территорию террасы левого берега реки Кумай в радиусе от 3 до 7 км, в восточной и северо-восточной части урочища Соколиные горы [1].

Комплекс состоит из нескольких групп разновременных памятников, компактно расположенных в восточной и юго-восточной части подножий горных вершин Акшокты, Ершокты и Жылынды. В карту памятников тюркского археолого-этнографического комплекса «Кумай» входят: тюркские ограды с изваяниями Косбатыр; конструкции эпохи бронзы; тюркские ограды с изваяниями Карагайлы (1-6); могильники эпохи бронзы Карагайлы (2-4); курганы, в том числе курган с «усами»; погребальные сооружения; курганная группа Балыкты-1; поселения и каменные изваяния (рисунок 1) [2,18].

рисунок 1. Косбатыр

На территории комплекса обнаружены каменные изваяния - балбалы, связанные с захоронениями. Они являются характерной чертой древнетюркской культуры и отражают особое отношение к культу предков и памяти о воинов. Балбалы комплекса «Кумай» представляют собой важный источник для изучения погребальных обрядов и сакральной организации пространства. Их наличие свидетельствует о высокой степени сакрализации территории и её значимости как центра коллективной памяти (рисунок 2) [2].

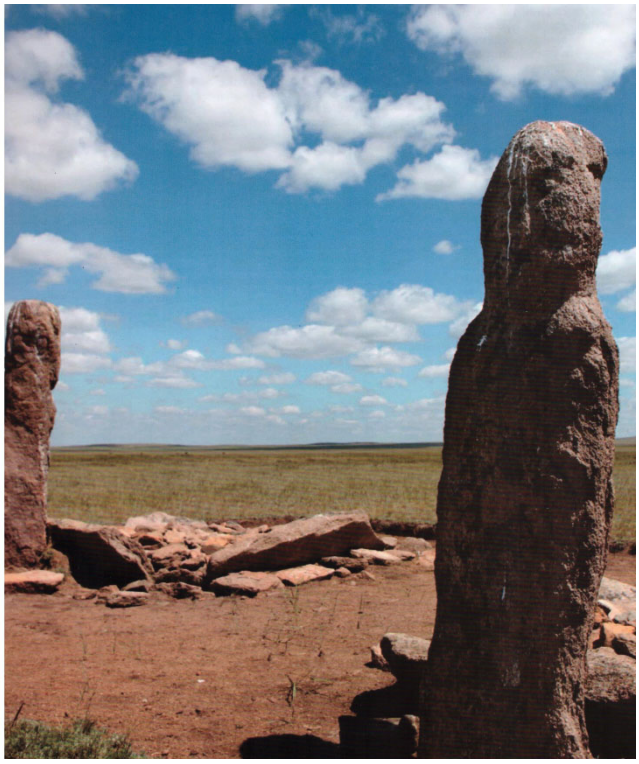


Рисунок 2. Балбалы комплекса Карагайлы 2

Анализ формы, расположения и символики каменных изваяний позволяет реконструировать представления древних тюрков о загробном мире, воинской чести и социальной иерархии. Мировоззрение тюркского общества характеризовалось идеей неразрывной связи человека и природы, что находило отражение в погребально-поминальной практике.

Особый интерес представляют находки из мужских и женских погребений, отличающиеся составом и назначением инвентаря. Они позволяют реконструировать социальный статус, род занятий и элементы повседневной жизни погребённых (рисунок 3). Так, обнаруженные в кургане Карагайлы-2 разноцветные «глазчатые» бусы свидетельствуют о сакральном и ритуальном назначении украшений. «Глазковый» узор, присутствующий в данных находках, и сегодня сохраняется в традиционном казахском быту как защитный символ, оберегающий от негативного воздействия (рисунок 4) [2,106,108].

Рисунок 3. Железное огниво и железный нож

Рисунок 4.Бусы

Комплекс «Кумай» представляет значительный интерес не только для археологов, но и для исследователей истории, этнографии и духовной культуры тюркских народов. Само название памятника отсылает к образу Кумай (Умай, Хумай) - одному из ключевых символов тюркской мифологии, связанному с идеями покровительства, продолжения рода и защиты жизни. Этот образ на протяжении веков занимал особое место в мировоззрении тюрков и находил отражение в обрядах, верованиях и материальной культуре, литературе. В мифологических представлениях тюркских народов данный образ близок к богине Умай и отражает представления о продолжении рода и сакральной защите.

Образ Хумай широко представлен в фольклорной и литературной традиции. В башкирском эпосе «Урал-батыр» она выступает как дочь небесных сил и помощница героя. В персидской литературе, в частности в поэме «Шахнаме», птица Хума символизирует счастье и царскую власть, а в поэзии Хафиза используется как метафора благополучия и высокой судьбы. Данный образ также рассматривается в трудах исследователей фольклора и этнографии, включая работы С. Аксакова, М. Ауэзова, С. Агиша и других авторов [3].

Комплекс «Кумай» представляет интерес не только для археологии, но и для смежных дисциплин - истории, этнографии и культурологии. Сопоставление археологических материалов с данными фольклора и мифологии способствует более глубокому пониманию культурных кодов тюркской цивилизации. Данный памятник, становится важным элементом общеевразийского исторического наследия.

Не менее значимым остаётся вопрос сохранения и популяризации комплекса «Кумай». Включение комплекса в научные исследования, образовательные и краеведческие программы способствует формированию исторического сознания и уважительного отношения к культурному наследию.

Таким образом, тюркский археолого-этнографический комплекс «Кумай», расположенный в Ерейментауском районе Акмолинской области, является значимым объектом историко-культурного наследия. Его изучение расширяет представления о развитии общетюркской цивилизации и подчёркивает роль региональных памятников в историческом процессе Казахстана.

Литература

1. Отчёт археологических исследований 2010 года . –Алматы .- 19с.
2. Досымбаева А., Нускабай А. «Тюркский археолого-этнографический комплекс Кумай» Астана, 2012.-146с.
3. Мифы народов мира/под ред. Токарева С. А. - М., Советская энциклопедия, 1992 г. - т.2 - 719 с.

Geographic Sciences

İqlim Dəyişikliyi Müstəvisində Qlobal Miqrasiya Prosesləri: Geosiyasi, Sosial-İqtisadi və Demoqrafik Transformasiyalar

Şəfəq Əliyeva Məmməd qızı
Azərbaycan Dövlət Pedaqoji Universitetinin Şəki filialı,müəllim

Açar Sözlər (Keywords)
İqlim miqrasiyası, ekoloji qaçqınlar, qlobal iqlim dəyişikliyi, Coğrafi İnformasiya Sistemləri (CİS), Aral dənizi böhranı, ekoloji determinizm, aridləşmə, yalançı urbanizasiya.
Climate migration, environmental refugees, global climate change, Geographic Information Systems (GIS), Aral Sea crisis, environmental determinism, aridization, pseudo-urbanization.

Xülasə

Antropogen fəaliyyət nəticəsində sürətlənən qlobal iqlim dəyişikliyi müasir coğrafiya elminin fundamental tədqiqat obyektinə çevrilmişdir. Bu proses sadəcə fiziki-coğrafi parametrlərin (temperatur, yağıntı balansı, dəniz səviyyəsi) dəyişməsi ilə məhdudlaşmır, həm də iqtisadi-sosial və demoqrafik coğrafiyanın ən ciddi böhranlarından birini — **iqlim miqrasiyasını** şərtləndirir. Bu akademik məqalədə ekoloji amillərin qlobal əhali məskunlaşmasına təsiri, miqrasiya arealları, regional risk ekosistemləri və prosesin geosiyasi fəsadları kompleks şəkildə təhlil olunur.

Giriş: Ekoloji Determinizm və Ekoloji Miqrasiyanın Teorik Çərçivəsi

Coğrafi determinizm və antropocoğrafiya nəzəriyyələri çərçivəsində insanın ətraf mühitlə qarşılıqlı əlaqəsi uzun illərdir öyrənilir. Lakin XXI əsrin ekoloji reallıqları bu əlaqəni krizis müstəvisinə daşımışdır. **İqlim miqrantları (və ya ekoloji qaçqınlar)** — ətraf mühitin qəfil və ya proqressiv pisləşməsi nəticəsində həyat şəraitinin kəskin dəyişməsi səbəbindən öz daimi yaşayış yerlərini müvəqqəti və ya daimi tərk etmək məcburiyyətində qalan şəxslərdir. (Biermann & Boas, 2010; Warner et al., 2009)

Klassik miqrasiya nəzəriyyələrində (məsələn, Ravenşteynin miqrasiya qanunları) "itələyici" (push) amillər sırasında iqtisadi və siyasi faktorlar öndə gəlirdisə, müasir dövrdə ekoloji itələyici amillər dominantlıq təşkil etməyə başlamışdır. (Black et al., 2011)

Qlobal İqlim Miqrasiyasının Coğrafi Determinantları

Ekoloji miqrasiyanı doğuran amillər coğrafi xarakterinə və təsir sürətinə görə iki əsas qrupa bölünür: (Warner et al., 2009)

A. Sürətli Baş Verən Kataklizmlər (Fast-onset events)

- **Hidrometeoroloji fəlakətlər:** Tropik siklonlar, tayfunlar, qəfil daşqınlar və sellər. (IPCC, 2022).
- **Demoqrafik effekt:** Əhalinin qısamüddətli, kütləvi və çox vaxt regional daxili yerdəyişməsinə səbəb olur. (UNHCR, 2023)

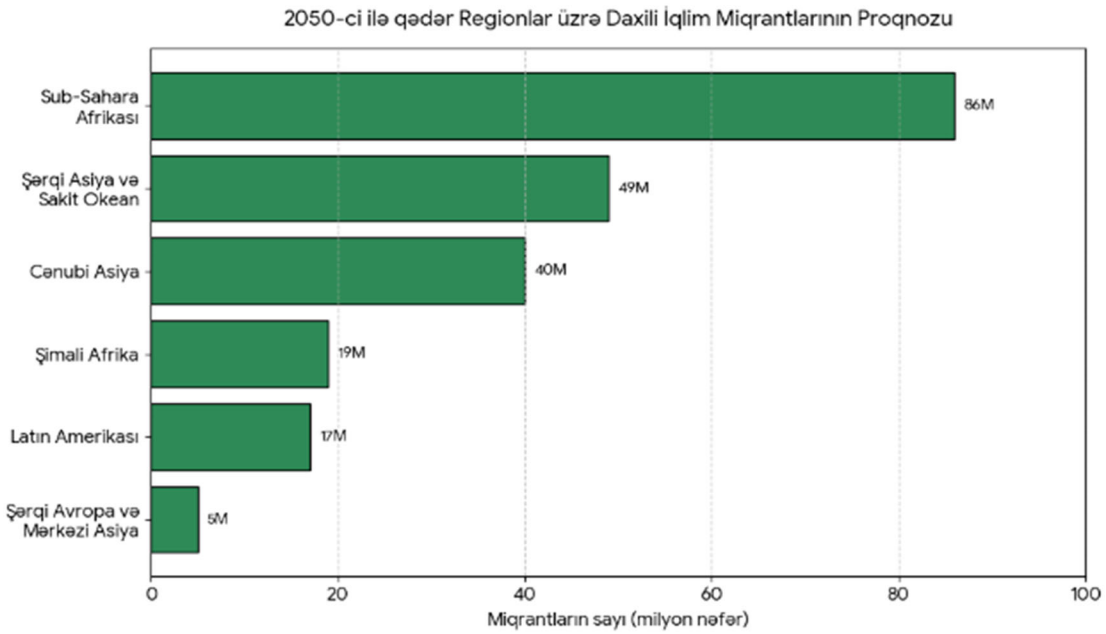
B. Tədricən İnkişaf Edən Proseslər (Slow-onset events)

- **Aridləşmə və səhralaşma:** Torpağın erroziyası, humid zonaların arid zonalara çevrilməsi və kənd təsərrüfatı resurslarının tükənməsi. (Myers, 2005)

- **Eustatik dəniz səviyyəsinin qalxması:** Dünya okeanının səviyyəsinin termal genişlənmə və qütb buzlaqlarının əriməsi hesabına ildə orta hesabla 3.4 mm artması. (IPCC, 2022)

Dünyanın Ən Həssas Regional Eko-Sistemləri və Miqrasiya Arealları

Beynəlxalq hesabatlara və coğrafi modelləşdirmələrə əsasən, iqlim miqrasiyasının episentrini təşkil edən coğrafi areallar aşağıdakı xəritə üzrə qruplaşdırılır: (Clement et al., 2021)



Sub-Sahara Afrikası (Sahel Zonası)

- **Coğrafi risk:** Seneqaldan Somaliyə qədər uzanan Sahel zonası səhralaşmanın ən aqressiv getdiyi regiondur. (Rigaud et al., 2018)
- **Nəticə:** Aqrar asılılığı olan əhali (çobanlar və fermerlər) otlaqların məhvi səbəbindən cənuba və sahil şəhərlərinə doğru hərəkət edir. Bu da yerli etnoslar arasında resurs müharibələrinə yol açır. (Clement et al., 2021).

Cənubi və Cənub-Şərqi Asiya (Delta Regionları)

- **Coğrafi risk:** Qanq-Brahmaputra (Banqladeş) və Mekonq deltaları dəniz səviyyəsinin qalxması və duzlaşma (salinizasiya) təhlükəsi altındadır. (Rigaud et al., 2018)
- **Nəticə:** Şirin su qatlarına dəniz suyunun sızması çəltikçilik təsərrüfatını tamamilə sıradan çıxarır. Banqladeşdə hər il minlərlə insan sahil zonalarından paytaxt Dəkkəyə miqrasiya edir. (Black et al., 2011).

Kiçik Ada Dövlətləri (SIDS)

- **Coğrafi risk:** Tuvalu, Kiribati, Maldiv və Marşall adalarının mütləq hündürlüyü dəniz səviyyəsindən cəmi bir neçə metr yüksəkdir. (Myers, 2002)

Bu dövlətlər üçün miqrasiya daxili deyil, tamamilə transmilli xarakter daşıyır. Bu da dünya tarixində ilk dəfə **"ərazisi olmayan dövlət"** və **"vətəndaşlığı olan, lakin coğrafi mkanı olmayan xalqlar"** fenomenini yaradır (Biermann & Boas, 2010).

Sosial-İqtisadi və Demoqrafik Transformasiyalar

İqlim miqrasiyası qlobal demoqrafik strukturu və urbanizasiya proseslərini kökündən dəyişir (Clement et al., 2021):

Göstərici	Əvvəlki Trendlər	İqlim Miqrasiyasının Təsiri
Urbanizasiya	Planlı iqtisadi mərkəzləşmə	"Yalançı urbanizasiya" və şəhərətrafı gecəqonduların (slums) idarəolunmaz böyüməsi
Əmək Bazarı	İxtisaslı kadr axını	Aqrar sektordan xidmət və tikinti sektoruna kütləvi, ixtisassız işçi qüvvəsi axını
Resurs İstehlakı	Balanslaşdırılmış təbii ehtiyat istifadəsi	Şəhər mərkəzlərində şirin su, kanalizasiya və enerji infrastrukturunun iflici

Geosiyasi Fəsadlar və Beynəlxalq Hüquq Boşluğu

Müasir siyasi coğrafiya və geosiyasət elmi iqlim miqrasiyasını qlobal təhlükəsizlik riski kimi nəzərdən keçirir (Myers, 2005). Ekoloji böhran sərhədləri tanımır, lakin siyasi sərhədlər miqrantların qarşısını kəsir.

- **Hüquqi Vakuum:** 1951-ci il Cenevrə Konvensiyasına əsasən, "qaçqın" statusu yalnız irqi, dini, milli və ya siyasi təqiblərə məruz qalan şəxslərə verilir. İqlim dəyişikliyi bu konvensiyada hüquqi əsas kimi tanınmır. Bu da iqlim qaçqınlarını beynəlxalq müdafiəsiz qoyur (Biermann & Boas, 2010).
- **Sərhəd Münaqişələri:** Banqladeşdən Hindistana, Şimali Afrikadan Avropa İttifaqına yönələn ekoloji miqrasiya dalğaları qəbul edən ölkələrdə sağ-populist siyasi cərəyanların güclənməsinə, ksenofobiyanın artmasına və sərhəd təhlükəsizliyinin militarizasiyasına səbəb olur (Myers, 2005; UNHCR, 2023

Nəticə və Adaptasiya Strategiyaları

İqlim miqrasiyası artıq uzaq gələcəyin ssenarisi deyil, müasir coğrafi reallıqdır. Bu qlobal böhranın coğrafi və sosial fəsadlarını minimuma endirmək üçün qlobal səviyyədə aşağıdakı strateji addımların atılması elmi dairələr tərəfindən tövsiyə olunur (Black et al., 2011; IPCC, 2022):

1. **"Yerində Adaptasiya" (In-situ adaptation):** Riski yüksək olan regionlarda dəniz bəndlərinin tikilməsi, quraqlığa davamlı transgen (GMO) kənd təsərrüfatı bitkilərinin tətbiqi və smart suvarma (damlama) sistemlərinin qurulması ilə əhalinin yerində qalmasını təmin etmək.
2. **Planlı Köçürmə (Planned Relocation):** Qaçılmaz fəlakət zonalarında olan əhalinin dövlət proqramları çərçivəsində daha təhlükəsiz coğrafi koordinatlara sistemli və humanist şəkildə köçürülməsi.
3. **Beynəlxalq Hüquqi Statusun Müəyyənləşdirilməsi:** BMT səviyyəsində iqlim miqrantlarının statusunu tənzimləyən xüsusi paktın qəbul edilməsi və qlobal kompensasiya fondlarının yaradılması (Clement et al., 2021).

Müasir coğrafiya elmi təbiət və cəmiyyət arasındakı bu disbalansı sadəcə qeydə almaqla kifayətlənməməli, həm də coğrafi informasiya sistemləri (CİS) vasitəsilə gələcək miqrasiya koridorlarını modelləşdirərək qlobal idarəetmə mexanizmlərinə elmi proqnozlar təqdim etməlidir.

Metodoloji İnnovasiya: Coğrafi İnformasiya Sistemləri (CİS) və Süni İntellektlə Miqrasiya Modelləşdirilməsi

Müasir coğrafiya elmində iqlim miqrasiyasını sadəcə baş verdikdən sonra tədqiq etmək kifayət deyil. **Proqnostik coğrafiya** çərçivəsində gələcək miqrasiya koridorlarının və "isti nöqtələrin" (hotspots) müəyyən edilməsi üçün CİS və Big Data (Böyük Verilənlər) texnologiyaları tətbiq olunur.

A. Peyk Monitorinqi və Ekoloji İndekslər

- **NDVI (Normallaşdırılmış Fərqli Vegetasiya İndeksi):** Peyk görüntüləri vasitəsilə Sahel zonası və ya Mərkəzi Asiyada yaşıl örtüyün azalması və səhralaşma sürəti ölçülür. Yaşıl örtüyün kritik həddə düşməsi həmin zonadan miqrasiyanın başlayacağına dair ilkin siqnaldır.
- **GRACE (Gravity Recovery and Climate Experiment):** Peyklər vasitəsilə yeraltı su hövzələrinin quruması izlənilir. Bu, xüsusilə regional su qıtlığı səbəbindən yaranacaq daxili köçləri öncədən proqnozlaşdırmağa imkan verir.

B. Süni İntellekt və Prediktiv Modelləşdirmə

- **Agent-Əsaslı Modelləşdirmə (Agent-Based Modeling - ABM):** Bu kompüter simulyasiyalarında hər bir fərd (və ya ailə) bir "agent" kimi kodlaşdırılır. Modelə temperatur artımı, məhsuldarlığın düşməsi və gəlir itkisi kimi faktorlar daxil edildikdə, süni intellekt əhalinin hansı faizinin, hansı zaman kəsiyində və hansı coğrafi istiqamətə (ən yaxın regional mərkəzə, yoxsa paytaxta) köçəcəyini yüksək dəqiqliklə hesablayır.

Keys Analizi (Case Study): Aral dənizi Hövzəsi və Mərkəzi Asiyanın Eko-Coğrafi Böhranı

Qlobal miqrasiya trendlərinin regional miqyasda regional dinamikasını anlamaq üçün Avrasiya materikinin ən həssas nöqtələrindən biri olan **Aral dənizi hövzəsi** (Özbəkistan, Türkmənistan, Qazaxıstan) klassik bir nümunədir.

Aral Dənizinin Quruması (Ekoloji fəlakət)

▼

Duz və Toksik Toz Fırtınaları (Aralqum səhrası)

▼

Kənd Təsərrüfatının İflası (Pambıqçılıq və aqrar böhran)

▼

Kütləvi Regional Miqrasiya (Nukus və Daşoğuzdan köçlər)

A. Ekoloji Tənəzzülün Mexanizmi

- Xətti antropogen müdaxilə (Amudərya və Sırdərya çaylarının suvarma üçün həddindən artıq istifadəsi) və qlobal istiləşmə nəticəsində Aral dənizi demək olar ki, tamamilə qurumuş və yerində **"Aralqum"** adlı yeni qum-duz səhrası yaranmışdır.
- Hər il bu quru dəniz dabanından milyonlarla ton zəhərli duz və pestisid qarışıq toz fırtınaları ilə ətraf rayonlara (Qaraqalpaqstan və Daşoğuz vilayəti) yayılır.

B. Demoqrafik Cavab reaksiyası (Miqrasiya Trayektoriyası)

1. **Mərhələ 1 (Yerli İflas):** Balıqçılıq və kənd təsərrüfatı tamamilə sıradan çıxdığı üçün ilkin olaraq sahilboyu yaşayış məntəqələri (məsələn, Muynak şəhəri) tamamilə boşalmışdır.
2. **Mərhələ 2 (Daxili və Transmilli Köç):** Ekosistemin korlanması və içməli suyun keyfiyyətinin düşməsi (bunun nəticəsində yaranan kəskin xəstəliklər) əhalini regional mərkəzlərə (Nukus), oradan isə iqtisadi cəhətdən daha stabil olan Daşkəndə və Rusiya Federasiyasına transmilli əmək miqrasiyasına məcbur etmişdir.
3. Aral keysi sübut edir ki, ekoloji miqrasiya çox vaxt özünü birbaşa deyil, **iqtisadi çətinlik pərdəsi altında** büruzə verir. İnsanlar rəsmi sənədlərdə "iş axtarmaq üçün" köçdüklərini qeyd etsələr də, bu işsizliyin fundamental kökündə ekoloji sistemin iflası dayanır (Black et al., 2011; Micklin, 2016).

Yekun Elmi Ümumiləşdirmə

XXI əsrin coğrafiyası artıq yalnız statik sərhədləri və relyef formalarını öyrənən fənn deyil. Müasir coğrafiya — dinamik, dəyişkən və insan-təbiət münasibətlərinin krizis nöqtələrini idarə edən strateji bir elmdir.

İqlim miqrasiyası təbiətin insana verdiyi ekoloji cavab reaksiyasıdır. Əgər qlobal emissiyalar azaldılmazsa və CİS əsaslı prediktiv regional adaptasiya proqramları tətbiq olunmazsa,

gələcəyin siyasi xəritəsi müharibələrlə deyil, iqlim miqrantlarının hərəkət trayektoriyaları ilə yenidən cızılacaqdır.

Ədəbiyyat Siyahısı (References)

1. **Biermann, F., & Boas, I. (2010).** Preparing for a warmer world: Towards a global governance system for climate refugees. *Global Environmental Politics*, 10(1), 60-88.
2. **Black, R., Bennett, S. R., Thomas, S. M., & Beddington, J. R. (2011).** Migration as a mitigation strategy for climate change. *Nature*, 478(7370), 447-449.
3. **Clement, V., Rigaud, K. K., de Sherbinin, A., Jones, B., Adamo, S., Schewe, J., ... & Shabahat, E. (2021).** *Groundswell Part 2: Acting on Internal Climate Migration*. World Bank, Washington, DC.
4. **IPCC (Intergovernmental Panel on Climate Change). (2022).** *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the IPCC. Cambridge University Press.
5. **Micklin, P. (2016).** The future Aral Sea: hope and despair. *Environmental Earth Sciences*, 75(9), 1-15.
6. **Myers, N. (2002).** Environmental refugees: a growing phenomenon of the 21st century. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1421), 609-613.
7. **Myers, N. (2005).** *Environmental Refugees: An Emergent Security Crisis*. 13th Economic Forum, Prague.
8. **Rigaud, K. K., de Sherbinin, A., Jones, B., Bergmann, J., Clement, V., Ober, K., ... & Midgley, A. (2018).** *Groundswell: Preparing for Internal Climate Migration*. World Bank, Washington, DC.
9. **UNHCR (United Nations High Commissioner for Refugees). (2023).** *Climate Change and Displacement*. UNHCR Official Reports and Strategic Frameworks.
10. **Warner, K., Ehrhart, C., de Sherbinin, A., Adamo, S., & Chai-Onn, T. (2009).** *In search of shelter: mapping the effects of climate change on human migration and displacement*. CARE International.

Sociological Sciences

Before the Weapons Sound: Media Power in Conflict

Eka Tkhilava, PhD
Caucasus International University; Doctor of Social Sciences
Maia Ekaladze
Doctoral Student, Caucasus International University

Abstract

This paper analyses the process of media involvement in the Georgian-Abkhazian conflict and its implications. There is a prevailing argument that disinformation holds greater power than tanks. This premise serves as the foundational point of departure for the present study — to examine the role of media in conflict coverage and to investigate how disinformation is capable of surpassing the force of a tank. The research led us to the origins of the Georgian-Abkhazian conflict; it is precisely these issues that are foregrounded, analysed, and interpreted throughout the paper.

Keywords: politics, media, Abkhazia, Georgia, disinformation.

Introduction

The Georgian-Abkhazian conflict of 1992-1993 was an artificially inspired action; its purpose was to detonate a geopolitical mine laid back during the Soviet Union. The Russian media served as an experienced ideological detonator in this process, thereby severing Abkhazia first from Georgia and then from Abkhazia itself. Consequently, the Abkhazian autonomy, which had been set on a separatist course for decades, found itself under the protectorate of Russia.

Our **research goal** is to study how a conflict becomes an observation zone and what dangers this process entails.

As our **research method**, we selected the qualitative historical and comparative approach; as research instruments, we employed case study analysis. We also outlined the political contexts. We used the "case study" method, which helped us better study and analyze the facts. Additionally, we utilised logical data analysis and comparison, which enabled us to identify similarities and differences between facts and events. Through content analysis, we were able to conduct an in-depth study of various document types, while deductive explanation models allowed us to draw logical conclusions based on existing views and facts.

The research showed the political-technological, ontological, and epistemological foundations of conflict coverage in the media and defined the empirical nature of the issue.

This paper is an attempt to show how the process of describing the conflict field developed, what gradations it underwent in modern media, and what challenges we faced.

While working on the paper, we utilised the study "The Vision Machine" by Italian sociologist Paul Virilio, which helped us view the phenomenon of conflict from a postmodern perspective.

"War teaches us to lose everything and to become what we were not..."

Thus thought Antoine de Saint-Exupéry.

It is perhaps difficult to explain the philosophy of war more precisely than this - losing everything and transforming into what we were not. And this transformation is by no means our best part, face, or ideal.

Virgil argued that even a war aimed at peace is a great evil.

Therefore, war is a symbol of loss, transformation, and great evil. This has been proven repeatedly to this day and will continue to be proven in the future.

It is also indisputable that the crime of war is not redeemed even by victory.

It is a Voltairean truth that "people who consider war not only an inevitable event, but useful and, consequently, desirable, are morally corrupt, dangerous, and terrible people."

There has never been a shortage of such people in the world.

The 1990s were no exception; These years proved chaotic and transformative for Georgia.

And when "the noise of weapons begins, laws fall silent" (Camus).

Laws were silent in Georgia in the 1990s as well.

Catching fish in muddy water turned out to be both simple and dangerous.

On 18 March 1989, a mass gathering of Abkhaz, numbering many thousands, took place in the village of Lykhny. Among those present was the party leadership of the Abkhazian Autonomous Republic as well; only the Chairman of the Council of Ministers, O. Zukhmaia, did not sign the "Lykhny Declaration."

This was, in effect, a denunciation of the Georgian people and of Georgia before "great, fraternal" Russia. Georgians were blamed for all the problems of the Abkhazian people. This was one of the principal documents through which the Abkhaz people demanded secession from the Georgian SSR, a demand that, in principle, contravened the constitutions of both the USSR and Georgia, which guaranteed the inviolability of the territorial integrity of the union republics.

But the water, as already mentioned, was muddy, the hidden noise of weapons was heard, and the law was silent.

Undoubtedly, conflict was inevitable -

Russian media tried to present Georgia as the starting point of everything. Russian journalist Vladimir Arsenyev wrote:

> For the mass of the Abkhaz people, who are asking for bread, it is not even known whether any will be brought at all. Dwellings without heating, without water, and often without electricity. An almost empty market, where one can only rarely buy a kilogram of potatoes at a price of up to 20 roubles. A curfew from 23:00 to 6:00; this is what I saw. And, moreover, checkpoints from the Krasnodar border all the way to western Georgia. Militia and internal troops armed with short-barreled assault rifles; in Sukhumi alone there are several such posts. (Vladimir Arsenyev, *Abkhazia without weapons, but without bread, gas, and electricity*, "Izvestia" newspaper, 21/03/1992, N 69).

Thus, several months before the start of the conflict, the ground was already being prepared so that the social situation created in Abkhazia would become the pretext for the outbreak of conflict.

In the 1990s, a deceptive calm was actually observed in Abkhazia; the echo of the civil war ongoing in Tbilisi was not felt beyond the Enguri River; state institutions were functioning; but the humanitarian and economic crisis was, of course, felt. However, this alone did not become the cause of the conflict; the problem was deeper and was rooted in the "Lykhny Declaration."

It was precisely the "Lykhny Declaration" that caused the tension in the political situation in Abkhazia, which laid the foundation for large-scale destabilisation. The Abkhazian sector of the Georgian National Movement also became active.

It is undoubtedly significant that Tbilisi, embroiled in civil war and social crisis, missed the processes ongoing in Abkhazia and the actions of Ardzinba's radical wing.

Accordingly, Ardzinba's separatist movement took advantage of this process and attempted to pass various political acts to separate from Georgia. For this very purpose, laws were adopted that were incompatible with the Constitution of Georgia; additionally, an Abkhazian

guard consolidated on a monoethnic basis was created, which in the future became the main armed formation.

Information about the Abkhazian war in the Russian media remains manipulative to this day and is aimed at covering the conflict with a pro-Russian narrative. The Russian media presents the problems of Abkhazia and Georgia from the following angles:

- Support for separatist forces: The Russian media covers the conflict as an Abkhaz national liberation movement, a struggle for Abkhazia's independence; Russia's own role is defined in terms of a "peacekeeping" function
- Disinformation and manipulation: Notably, the Russian media ignores the war crimes and facts of ethnic cleansing committed by the Abkhazians, in which the Russian army itself directly participated.
- Discrediting Georgia: The Russian media accuses the Georgian government of pursuing an "aggressive policy" and provoking hostilities.

The fact is that the Russian media is trying to rewrite the history of this war and covers up Russia's participation in the 1992-1993 confrontation. In this way, the Russian media attempts to justify the occupation of Abkhazia.

Therefore, the conflict in Abkhazia (1992-1993) started much earlier than the Georgian armed forces entered there to establish order on the railway on August 14, 1992. However, Georgia entered a war it was doomed to lose from the very beginning.

A curious symbolism also attends the armed conflict in Abkhazia, which lasted thirteen months and thirteen days; according to unverified data, more than 13,000 people lost their lives in the conflict. Thus the number thirteen proved fateful for Georgia as well. It is no coincidence that in France there are to this day no houses bearing the number 13; in the United States there are no buses numbered 13; many hotels have no room number thirteen; trains have no thirteenth carriage; and ships have no thirteenth cabin…Georgia, too, became a kind of symbolic prisoner of the number thirteen.

In the Abkhazian conflict, more than 1,000 persons were also declared missing without trace. Up to 300,000 people became internally displaced. Almost thirty-four years after the outbreak of the war, the conflict remains "frozen," and the status quo is unchanged.

It is noteworthy and significant that the Supreme Council of the Autonomous Republic of Abkhazia has documented the stories of 4,321 people who were tortured and shot by Russia during and after the occupation of Abkhazia.

Today, the interests underlying Russian policy are becoming even clearer: it has grown more aggressive and seeks to bring Abkhazia's economic resources fully under its control. Russia's interests also extend to Abkhazia's strategic assets.

The Abkhaz population is evidently dissatisfied, having expected different outcomes from the conflict. The language of Russian policy, however, is the language of blackmail: it threatens to suspend budgetary funding for Abkhazia's social and economic sectors.

In 2024, the meeting between the de facto president of Abkhazia, Aslan Bzhania, and the Deputy Chief of Staff of the Russian Presidential Administration, Dmitry Kozak, was intensively covered on Abkhazian opposition Telegram channels. Discussion and excitement were triggered by the protocol adopted within the framework of this meeting, according to which Russia would suspend funding to Abkhazia if the agreement on the recognition and enforcement of court and arbitration decisions on economic matters were not ratified.

Additionally, Russia demanded ensuring the entry into force of the agreement on the implementation of investment projects on the territory of Abkhazia for Russian legal entities and lifting restrictions on the commercial real estate market for Russian investors. Therefore, funding would occur only and exclusively if Abkhazia fully implemented Russian narratives on its territory.

Abkhazia remains in a profound impasse, with no way out in sight. It is compelled to accept every Russian condition, because its economic situation depends entirely on Russia. Abkhazia receives more than 3.9 billion rubles annually from Russia, which covers wages, pensions, and other essential social programmes. In case of resistance, Abkhazia faces complete economic stagnation; therefore, it follows the principle that "it is better to go along than to be dragged in" and fulfils all of Russia's demands completely.

The resistance created from time to time within Abkhazia occurs not with the demand for separation and liberation from Russia, but rather as a counterweight to "better adjust" to Russia. However, despite this, funding from Russia has been on a declining trend in recent years; the occupying country is trying to create an even more severe crisis for the region to finally lead it into a dead end and make it completely dependent on it economically.

Thus, the situation of the Abkhazian population is extremely difficult. Although the Abkhaz opposition and civil society sector have made vocal statements criticising the de facto government for openly serving Russian interests, the situation grows more difficult by the day. Relations between Abkhazia and Russia not infrequently reach an impasse. For part of the population it is becoming clear that they are being denied any opportunity for independent development, and that Russian ultimatums pose what amount to existential threats.

The Russian media respond actively to every such instance of resistance and aggravate the situation further, signalling what fate awaits Abkhazia should it resist Russia, and what economic stagnation will mean for the Abkhaz population.

Consequently, the Russian media continue to pursue their long-standing and consistent policy. Russia is presented as the embodiment of the "big brother" ideal, while every neighbouring state, Georgia included, is portrayed as ill-disposed towards it.

Everyday reality demonstrates how difficult the geopolitical situation is for the countries neighbouring Russia as well. A dictatorial state poses a threat not only within its own borders, for under a dictatorial regime everyone is afraid, and everyone is afraid even of themselves.

Ultimately, every one of these processes can be explained by the well-known phrase of the Russian Empress Catherine II: "Politics is not a hospital! Those who are weak are the first to be dragged out!"

There is only one way out: either we must not be weak, or dictatorial states must not be allowed to see our weakness.

Peace must be safeguarded before weapons speak and the law falls silent.

Bibliography

1. Vladimir Arsenyev, *Abkhazia without weapons, but without bread, gas, and electricity*, "Izvestia" newspaper, 21/03/1992, No. 69
2. https://socialjustice.org.ge/ka/products/ruseti-afkhazettan-ultimatumis-enit-itsqebs-laparaks.
3. https://exclusivenews.ge/post/562585.
4. https://www.mediachecker.ge/mediagaremo/article/94364-media-okupirebul-afkhazethshi.
5. https://mtisambebi.ge/news/item/1495-rogor-axamebda-da-xvretda-ruseti-qartvelebs-apxazetis-omshi.
6. https://www.spekali.tsu.ge/index.php/ge/article/viewArticle/16/283.
7. https://www.ghn.ge/news/183625-rogor-ratom-da-ra-miznit-daitsqo-omi-afkhazetshi-25-tslis-tsin.
8. https://gnomonwise.org/ge/publications/policy-papers/103.

Publisher.agency: Proceedings of the 14th International Scientific Conference «Foundations and Trends in Research» (September 10-11, 2026). Copenhagen, Denmark, 2026. 274p

editor@publisher.agency

https://publisher.agency

University of Copenhagen

Lindestien, 23

2605 Brøndbyøster Copenhagen,

Denmark