

УДК 004.657:61

[https://doi.org/10.52058/2786-4952-2025-5\(51\)-2802-2815](https://doi.org/10.52058/2786-4952-2025-5(51)-2802-2815)

Зінченко Наталія Олександрівна кандидат філософських наук, доцент, доцент кафедри філософії та суспільних наук, Полтавський державний медичний університет, м. Полтава, <https://orcid.org/0000-0002-3247-3836>

Абрамчук Ольга Миколаївна кандидат біологічних наук, доцент кафедри фізіології людини і тварин, Волинський національний університет імені Лесі Українки, м. Луцьк, <https://orcid.org/0000-0002-9479-2111>

Кузик Петро Васильович кандидат медичних наук, доцент кафедри патологічної анатомії, Національний медичний університет імені О.О. Богомольця, м. Київ, <https://orcid.org/0000-0001-9352-4513>

ЗАСТОСУВАННЯ ВЕЛИКИХ ДАНИХ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ БІОМЕДИЧНИХ ДОСЛІДЖЕНЬ.

Анотація. Досліджено роль великих даних як стратегічного ресурсу в трансформації біомедичних досліджень на сучасному етапі розвитку медичної науки. Актуальність зумовлена зростанням обсягів і складністю даних, що генеруються внаслідок упровадження високопродуктивних технологій секвенування, цифрової візуалізації, електронних медичних записів і сенсорного моніторингу, а також потребою в підвищенні точності, швидкості та адаптивності клінічних рішень. Метою дослідження є обґрунтування значущості великих даних для підвищення результативності біомедичних досліджень шляхом оцінки сучасних технологічних рішень, виявлення їхніх переваг і обмежень, а також формулювання науково обґрунтованих рекомендацій щодо оптимізації застосування. Методологічну основу становить поєднання структурно-функціонального, порівняльного, інформаційно-аналітичного та системного підходів, що уможливило здійснення комплексної оцінки технічних, етичних та організаційних аспектів застосування великих даних в біомедичній сфері. Встановлено, що великі дані забезпечують перехід до індуктивних моделей аналізу, інтеграції реальних клінічних даних (real-world evidence), підвищують релевантність доказової бази та сприяють персоналізації медицини. Виявлено основні обмеження: фрагментарність даних, низький рівень стандартизації, недостатня прозорість алгоритмів та дефіцит етичного супроводу. Доведено необхідність упровадження стандартів HL7 FHIR, застосування explainable AI, створення міждисциплінарних команд і оновлення нормативного середовища для гарантування безпечної й ефективної інтеграції аналітики в медичну практику. Наукова новизна полягає

в систематизації парадигмальних зрушень у біомедичних дослідженнях під впливом великих даних і розробці концептуальних рекомендацій для подолання бар'єрів їхнього застосування. Перспективами подальших досліджень визначено розробку індикаторів якості медичних великих даних, етичних моделей управління алгоритмізованими процесами та цифрових екосистем для підтримки прийняття клінічних рішень.

Ключові слова: алгоритмічний аналіз, інтероперабельність, доказова база, персоналізована медицина, етичні ризики, штучний інтелект, цифрова трансформація.

Zinchenko Nataliia Oleksandrivna PhD in Philosophy, Associate Professor, Associate Professor of the Department of Philosophy and Social Sciences, Poltava State Medical University, Poltava, <https://orcid.org/0000-0002-3247-3836>

Abramchuk Olha Mykolayivna PhD in Biology, Assistant Professor of the Department of Human and Animal Physiology, Lesya Ukrainka Volyn National University, Lutsk, <https://orcid.org/0000-0002-9479-2111>

Kuzyk Petro Vasylovych PhD in Medical Sciences, Docent, Associate Professor of the Department of Pathological Anatomy, Bogomolets National Medical University, Kyiv, <https://orcid.org/0000-0001-9352-4513>

USING BIG DATA TO ENHANCE THE EFFICIENCY OF BIOMEDICAL RESEARCH

Abstract. The study explores the role of big data as a strategic resource in transforming biomedical research at the current stage of scientific development. The topic's relevance stems from the exponential growth in the volume and complexity of data generated through high-throughput sequencing technologies, digital imaging, electronic health records, biosensor monitoring, and the growing need for improved accuracy, speed, and adaptability in clinical decision-making. The aim of the study is to substantiate the significance of big data in enhancing the efficiency of biomedical research by evaluating current technological solutions, identifying their advantages and limitations, and formulating evidence-based recommendations for optimizing their application. The methodological framework combines structural-functional, comparative, analytical, and systems-based approaches, enabling a comprehensive assessment of big data implementation's technical, ethical, and regulatory aspects in the biomedical field. It has been established that big data enables a shift toward inductive models of analysis, integration of real-world clinical data, improved relevance of evidence bases, and the advancement of personalized medicine. Key challenges have been identified, including data fragmentation, lack of interoperability, limited standardization, insufficient algorithmic transparency,

and ethical concerns regarding data privacy and decision-making responsibility. The need for adopting HL7 FHIR standards, promoting explainable AI, forming interdisciplinary research teams, and updating legal frameworks for secure and effective data integration has been demonstrated. The scientific novelty lies in the systematization of paradigm shifts in biomedical research driven by big data and the development of conceptual recommendations to overcome implementation barriers. Future research prospects include the development of quality indicators for medical big data, ethical governance models for algorithmic decision-making, and the design of adaptive digital ecosystems to support clinical decision processes.

Keywords: algorithmic analysis, interoperability, evidence base, personalized medicine, ethical risks, artificial intelligence, digital transformation

Постановка проблеми. Сучасна біомедична наука розвивається в умовах стрімкого зростання обсягів даних, що формуються внаслідок широкого застосування високопродуктивних технологій секвенування, цифрової візуалізації, електронних медичних записів та біосенсорного моніторингу. Зростання масштабів і складності цих інформаційних потоків висуває нові вимоги до методів обробки, інтеграції та аналізу, що актуалізує необхідність пошуку ефективних рішень для перетворення великих даних на практично значущі результати. Відсутність уніфікованих інструментів обробки, складність забезпечення міждисциплінарної інтероперабельності та обмежений доступ до структурованих масивів даних значно стримують науковий прогрес у галузі.

Застосування концепцій і методів роботи з великими даними (Big Data) безпосередньо пов'язане з вирішенням актуальних наукових і практичних завдань, таких як розробка персоналізованих терапевтичних стратегій, прогнозування перебігу патологічних станів, оптимізація клінічних рішень і скорочення часу трансляції результатів фундаментальних досліджень у медичну практику. Інтеграція великих даних у біомедичні дослідження відкриває нові можливості для побудови точних моделей біологічних процесів, підвищення надійності доказової бази та забезпечення стійкого наукового розвитку.

Аналіз останніх досліджень і публікацій. Огляд наукових праць щодо використання великих даних для підвищення ефективності біомедичних досліджень дає можливість виокремити чотири основні змістові напрями, які репрезентують сучасні підходи до оцінки потенціалу, технологій, ризиків та практик їхнього впровадження в біомедицині.

Перший напрям зосереджено на загальних тенденціях і потенціалі великих даних для розвитку біомедичних досліджень. М. Мойсеєнко, М. Кузишин, Л. Туровська, Ю. Мазуренко, М. Петришин та О. Мазуренко розглядають можливості великих мовних моделей штучного інтелекту в медицині, акцентуючи на викликах точності та прозорості їхнього застосування [1].

С. Кремін (С. Cremin), С. Деш (S. Dash) і Х. Хуанг (X. Huang) аналізують історичні аспекти та сучасні тенденції впровадження великих даних, наголошуючи на їхньому значенні для формування нових клінічних моделей [2]. Х. Янг (X. Yang), К. Хуан (K. Huang), Д. Янг (D. Yang), В. Чжао (W. Zhao) та Х. Чжоу (X. Zhou) досліджують вплив великих даних на розвиток персоналізованої медицини, визначаючи переваги й обмеження такого підходу [3]. Й. Пей (Y. Pei) та Дж. Янг (J. Yang) підкреслюють важливість інтеграції штучного інтелекту з великими даними в клінічних застосуваннях для підвищення точності діагностичних і терапевтичних рішень [4]. Подальші дослідження цього напрямку доцільно зосередити на конкретизації клінічних сценаріїв та економічній оцінці впровадження великих даних у сферу охорони здоров'я.

Другий напрям охоплює технології обробки великих даних, зокрема алгоритми машинного навчання та інтегративні підходи. К. Таха (K. Taha) проводить комплексний емпіричний аналіз алгоритмів машинного навчання, визначаючи їхню практичну ефективність у медичних дослідженнях [5]. А. Рехман (A. Rehman), С. Наз (S. Naz) та І. Раззак (I. Razzak) досліджують актуальні підходи та технологічні рішення для аналізу великих даних у медичних закладах, демонструючи їхню ефективність для підвищення якості медичних послуг [6]. Д. Саргіотіс (D. Sargiotis) акцентує на застосуванні штучного інтелекту й великих даних для прискорення біомедичних відкриттів та оптимізації процесів досліджень [7]. Б. Мірза (B. Mirza), В. Ванг (W. Wang), Дж. Ванг (J. Wang), Х. Чхой (H. Choi), Н. С. Чунг (N. C. Chung) та П. Пін (P. Ping) презентують інтегративні підходи аналізу великих біомедичних даних, що уможливорює виявлення нових взаємозв'язків та патернів у медичній інформації [8]. Майбутні дослідження варто сфокусувати на вдосконаленні алгоритмів із високою обчислювальною ефективністю та їхню адаптацію до роботи в реальному часі.

Третій напрям розглядає управлінські й організаційні аспекти впровадження технологій великих даних у медичні дослідження та охорону здоров'я. С. Даш (S. Dash), С. К. Шакаявар (S. K. Shakyawar), М. Шарма (M. Sharma) та С. Каушик (S. Kaushik) аналізують сучасні тенденції управління великими даними, визначаючи організаційні й технічні умови ефективного впровадження технологій [9]. А. Бенерджі (A. Banerjee), Ч. Чакраборті (C. Chakraborty), А. Кумар (A. Kumar) та Д. Бісвас (D. Biswas) наголошують на значенні IoT та великих даних для розвитку сучасних медичних технологій і платформ моніторингу [10]. Ф. Денг (F. Deng), Дж. Хуанг (J. Huang), Х. Юань (X. Yuan), Ц. Ченг (C. Cheng) та Л. Чжан (L. Zhang) аналізують ефективність різних алгоритмів машинного навчання для аналізу структурованих біомедичних даних, що є важливим аспектом організаційних рішень [11]. Б. Вонг (B. Wang) та Л. Лі (L. Li) досліджують технологічні аспекти управління біомедичними великими даними, підкреслюючи значення безпеки та стандар-

тизації управлінських процесів [12]. У подальших роботах важливо зосередитись на розвитку моделей інтегрованого управління великими даними та створенні міжнародних стандартів щодо обміну й зберігання медичної інформації.

Четвертий напрям зосереджений на вивченні етичних, правових та регуляторних проблем, пов'язаних із застосуванням великих даних у біомедичних дослідженнях. А. Ферретті (A. Ferretti), М. Ієнка (M. Ienca), С. Херст (S. Hurst) і Е. Ваєна (E. Vayena) визначають нові виклики для етичних комітетів, пов'язані з обробкою великих біомедичних даних, наголошуючи на необхідності посилення механізмів захисту конфіденційності [13]. В. Навале (V. Navale), Д. ван Капплер (D. von Kaeppler) та М. МакОліфф (M. McAuliffe) аналізують біомедичні платформи управління даними, акцентуючи на важливості етичного та безпечного обігу медичної інформації [14]. Завершує цей напрям праця А. Банерджі (A. Banerjee), С. Чакраборті (S. Chakraborty), А. Кумар (A. Kumar) та Д. Бісвас (D. Biswas), що підкреслює актуальність упровадження сучасних технологій для збереження безпеки медичних даних і відповідності міжнародним стандартам [15]. Перспективним напрямом для наступних досліджень є розроблення узгоджених міжнародних політик і правових стандартів для регулювання біомедичних досліджень на основі великих даних.

Таким чином, аналіз наукових праць засвідчує, що ефективність застосування великих даних у біомедичних дослідженнях зумовлене чітким визначенням потенціалу цих технологій, розробкою надійних алгоритмів машинного навчання, ефективним організаційним управлінням великими обсягами даних та розв'язанням етичних і правових проблем щодо конфіденційності й регулювання їхнього використання.

Мета і завдання статті. Метою дослідження є обґрунтування ролі великих даних у підвищенні ефективності біомедичних досліджень шляхом аналізу сучасних технологічних підходів, оцінки їхньої практичної результативності та формування рекомендацій щодо оптимізації застосування у цій сфері.

Для досягнення мети визначено такі завдання:

1. Систематизувати сучасні підходи до застосування великих даних у біомедичній науці та визначити їхню роль у зміні аналітичних моделей і парадигм доказової медицини.

2. Проаналізувати основні інструменти й технології обробки великих даних у біомедичних дослідженнях з урахуванням їхньої алгоритмічної складності, інтероперабельності та впливу на процеси прийняття клінічних рішень.

3. Виявити основні проблеми впровадження великих даних у біомедичну практику та розробити рекомендації для вдосконалення підходів до їхнього збору, обробки й використання з урахуванням технічних, етичних та нормативних викликів.

Виклад основного матеріалу. Застосування великих даних у біомедичній науці поступово трансформується з експериментального інструмента в

системний компонент дослідницьких і клінічних процесів. У межах сучасних наукових підходів значні обсяги медичних даних розглядаються не лише як джерело емпіричної інформації, а як основа для формування нових гіпотез, перевірки взаємозв'язків між біомедичними параметрами та створення предиктивних моделей. У системі доказової медицини великі дані сприяють переходу від традиційних рандомізованих досліджень до аналізу реальної інформації про пацієнтів, що забезпечує оцінювання ефективності втручань у гетерогенному середовищі. Це розширює потенціал персоналізованої медицини, підвищує швидкість формування доказової бази та зменшує вартість дослідницьких процедур (табл. 1).

Таблиця 1

Підходи до застосування великих даних у біомедичних дослідженнях та їхня інтеграція в систему доказової медицини

Підхід до застосування великих даних	Основна характеристика	Зв'язок із доказовою медициною
Індуктивне моделювання	Побудова висновків на основі аналізу великих масивів даних без попередніх гіпотез	Сприяє формуванню нових клінічних гіпотез із високим рівнем статистичної обґрунтованості
Реєстрові дослідження	Використання даних електронних медичних записів та національних реєстрів	Дає можливість аналізувати ефективність методів лікування в реальних клінічних умовах
Персоналізований аналіз	Обробка даних конкретного пацієнта з урахуванням геноміки, біомаркерів та поведінкових чинників	Забезпечує точність клінічних рішень у межах доказової моделі
Системна інтеграція даних	Об'єднання різних типів біомедичних даних (оміка, зображення, записи, сенсори)	Сприяє формуванню комплексних доказів про ефективність та безпеку втручань
Машинне навчання для аналізу ефектів	Застосування алгоритмів для виявлення закономірностей у клінічних даних	Підвищує швидкість аналізу доказів і знижує ризик людських помилок

Джерело: сформовано авторами на основі [2; 3; 4; 5; 6; 7; 8; 9; 13]

Застосування індуктивного моделювання сприяє формулюванню нових наукових гіпотез на основі статистично обґрунтованих залежностей, виявлених у гетерогенних масивах даних. Це особливо актуально для ситуацій, де традиційні експериментальні методи не забезпечують необхідного обсягу доказів. Реєстрові дослідження на основі електронних медичних записів, демонструють високу ефективність у вивченні наслідків лікування, у тому числі рідкісних побічних ефектів або варіабельності відповіді на терапію в різних популяційних групах.

Персоналізований аналіз даних у поєднанні з геномною інформацією та біомаркерами дає змогу адаптувати клінічні стратегії до індивідуальних

характеристик пацієнта, що відповідає пріоритетам доказової медицини третього покоління. Системна інтеграція різнорідних даних уможливорює отримання цілісного уявлення про патологічні процеси та ефективність втручань, знижуючи ймовірність помилкових висновків через фрагментарність інформації. Застосування машинного навчання для аналізу ефектів сприяє підвищенню продуктивності обробки даних та виявленню слабких, але стабільних сигналів, які в традиційній статистиці залишаються поза увагою. У практичних умовах охорони здоров'я ці підходи реалізуються як клінічні рішення, що ґрунтуються на реальних даних пацієнтів, автоматизованій аналітиці та динамічному оновленні доказової бази, що значно розширює можливості медичної науки й практики.

Інструменти та технології обробки великих даних у біомедичних дослідженнях розвиваються відповідно до високих вимог до точності, швидкодії та масштабованості аналітичних систем. Сучасні біомедичні дані є різнорідними за своєю структурою, джерелами та форматами – від геномних послідовностей і зображень до показників фізіологічного моніторингу та клінічних записів. Це потребує застосування високопродуктивних технологій, здатних обробляти великі обсяги інформації в реальному або майже реальному часі. Алгоритмічна складність таких інструментів визначає їхню здатність ефективно аналізувати складні патерни, тоді як інтероперабельність забезпечує можливість інтеграції різнорідних джерел даних в єдине дослідницьке або клінічне середовище (табл. 2).

Таблиця 2

**Характеристика інструментів обробки великих даних у
біомедичних дослідженнях**

Інструмент або технологія	Алгоритмічна складність	Інтероперабельність	Основне біомедичне застосування
Apache Spark	Висока ефективність для розподіленої обробки даних, масштабована обчислювальна модель	Підтримка багатьох джерел даних (JSON, CSV, HL7, FHIR)	Аналіз великих реєстрів, предиктивне моделювання захворювань
TensorFlow	Потужні алгоритми глибокого навчання, складність залежить від архітектури моделі	Інтеграція з клінічними БД, PACS, біомедичними API	Автоматична обробка медичних зображень, діагностичні моделі
Hadoop	Масштабована обробка неструктурованих даних, помірна складність	Працює з різноманітними форматами даних	Архівування, очищення та агрегація клінічних даних
KNIME	Гнучка графічна побудова аналітичних потоків, низька порогова складність	Плагіни для інтеграції з базами біологічних і клінічних даних	Візуалізація та початковий аналіз даних з біобанків
BioConductor	Статистичні методи для аналізу омічних даних, складність залежить від сценарію	Інтеграція з іншими R-бібліотеками та омічними платформами	Аналіз експресії генів, епігенетичних змін, мутацій

Джерело: сформовано авторами на основі [1; 5; 10]

Наведені в табл. 2 інструменти відображають різнопрофільні рішення, що сприяють адаптації аналітичних процесів до складності біомедичних даних та забезпечення їхньої сумісності з різними джерелами й форматами. Apache Spark [16] як розподілена обчислювальна платформа дає можливість паралельно обробляти великі реєстрові масиви медичних записів, зокрема у форматах HL7 (Health Level Seven) [17] та FHIR (Fast Healthcare Interoperability Resources) [18], що використовуються в лікарняних інформаційних системах. TensorFlow, розроблений компанією Google [19], забезпечує побудову складних моделей глибокого навчання, зокрема для автоматизованого розпізнавання медичних зображень у системах PACS (Picture Archiving and Communication System) [20], які є стандартом у цифровій радіології.

Hadoop, технологія обробки неструктурованих даних у розподіленому середовищі, сприяє ефективному архівуванню та агрегуванню клінічних даних із різних систем, не потребуючи попередньої стандартизації. Його використання є визначальним на етапах підготовки даних до аналітики. KNIME [21] (Konstanz Information Miner) дає змогу будувати аналітичні сценарії за допомогою візуального інтерфейсу, що особливо важливо для біомедичних дослідників, які не володіють програмуванням, але працюють з даними біобанків і клінічних лабораторій. BioConductor [22] – це набір програмних пакетів у середовищі R, орієнтованих на біоінформатичний аналіз, що широко застосовується для обробки омічних даних, зокрема транскриптоміки, метиломіки та варіантного аналізу.

У реальних умовах ці інструменти комбінуються в аналітичні пайплайни, що охоплюють попереднє очищення, нормалізацію, статистичне тестування, побудову предиктивних моделей та візуалізацію результатів. Ефективність їхнього застосування залежить не лише від технічних характеристик, а й від здатності дослідницьких установ інтегрувати дані з різних джерел та забезпечити міждисциплінарну взаємодію [8]. В умовах сучасної біомедичної науки саме інтеоперабельність і масштабованість цих платформ є визначальними чинниками в прискоренні відкриттів і трансляції результатів досліджень у клінічну практику.

Упровадження технологій великих даних істотно впливає на зміну парадигм у біомедичних дослідженнях, зокрема щодо методів аналізу інформації та процесів прийняття клінічних рішень. Традиційна модель наукового дослідження ґрунтувалася на перевірці гіпотез із використанням обмежених вибірок, тоді як сучасні підходи орієнтовані на обробку повноти наявних даних – від омічних профілів до реєстрових спостережень – у реальному часі. Відбувається перехід від детерміністичних до статистично-прогностичних моделей, що ґрунтуються на машинному навчанні, глибоких нейронних мережах та багатошаровому аналізі складних взаємозв'язків. Це дає можливість формувати точні, адаптивні та персоналізовані клінічні рішення, що враховують індивідуальні характеристики пацієнтів та контекстуальні чинники середовища (табл. 3).

Таблиця 3

Парадигмальні зрушення в біомедичних дослідженнях під впливом технологій великих даних

Парадигмальний зсув	Традиційний підхід	Новий підхід під впливом великих даних	Приклад застосування
Джерело даних	Обмежена вибірка з контролем за змінними	Великі обсяги гетерогенних даних у реальному часі	Застосування EHR (Electronic Health Records) для досліджень ефективності ліків
Аналітична модель	Регресійні або кореляційні моделі	Алгоритми машинного навчання, глибокі нейронні мережі	Прогнозування ризику інсульту на основі багатофакторного аналізу
Формат доказів	Результати рандомізованих контрольованих досліджень	Індуктивно сформовані висновки на основі реальних даних	Розробка клінічних рекомендацій із застосуванням реальних клінічних даних
Характер рішення	Статичне, стандартизоване	Динамічне, персоналізоване	Модифікація терапії на основі індивідуального геномного профілю
Роль лікаря / дослідника	Суб'єктивна інтерпретація результатів	Підтримка рішень на основі автоматизованого аналізу	Застосування CDS (Clinical Decision Support) систем у щоденній практиці

Джерело: сформовано авторами на основі [2; 3; 5; 8; 9; 12; 13]

Зміни, наведені в таблиці, на практиці реалізуються через поступову інтеграцію аналітичних платформ у робочі середовища медичних установ. Наприклад, використання електронних медичних записів (EHR) для побудови предиктивних моделей у системах CDS (Clinical Decision Support) дає змогу автоматизувати виявлення клінічних ризиків, таких як передчасна серцева недостатність чи діабет другого типу, ще до прояву симптомів [23], [24]. При цьому рішення ґрунтуються на безперервному аналізі нових даних, що надходять у режимі реального часу. Алгоритми машинного навчання адаптуються до особливостей пацієнтів і виявляють нетипові клінічні патерни, що залишаються поза увагою в межах традиційного клінічного мислення [25]. У сфері досліджень це забезпечує можливість швидкого оновлення доказової бази з урахуванням нових масивів реальних клінічних даних, що істотно скорочує шлях від експерименту до клінічної практики [26]. Такий підхід сприяє переходу до медицини від лікування до превенції завдяки проактивному аналізу індивідуального ризику та точній стратифікації пацієнтів [27].

Попри значні досягнення у сфері аналітики великих даних, їхнє широке впровадження в біомедичну практику супроводжується низкою комплексних проблем, що мають як системний, так і прикладний характер. Зокрема, на рівні технічної реалізації досі існують перешкоди, зумовлені фрагментацією даних, їхньою недостатньою структурованістю, а також несумісністю форматів, що застосовуються в медичних закладах. Це ускладнює інтеграцію даних із різних джерел – електронних медичних записів, лабораторних систем, сенсорів або геномних баз – і обмежує потенціал побудови комплексних моделей. Окрему проблему становить забезпечення інтероперабельності між платформами різних виробників, особливо у випадках транснаціонального обміну клінічно значущою інформацією [6].

Не менш важливими є етичні виклики, зумовлені питаннями конфіденційності, згоди на обробку персональних медичних даних, а також ризиками дискримінації в разі некоректної інтерпретації результатів алгоритмів. У деяких випадках існує реальна загроза того, що автоматизовані рішення, побудовані на статистичній закономірності, можуть суперечити інтересам конкретного пацієнта, особливо якщо модель не пояснює свої дії або функціонує як «чорна скринька» [11]. Потреба в етичному супроводі технологій великих даних стає все актуальнішою, зважаючи на зростання кількості цифрових платформ для моніторингу здоров'я та дистанційної медицини.

Окрему категорію становлять регуляторні обмеження, що охоплюють як питання правового статусу аналітичних систем на базі великих даних, так і процедури сертифікації, перевірки точності, репрезентативності та безпечності використання таких систем у клінічній практиці [14]. У багатьох юрисдикціях законодавство не встигає за розвитком технологій, що створює правовий вакуум щодо відповідальності за помилки алгоритмів або порушення прав суб'єктів даних.

З урахуванням виявлених тенденцій і бар'єрів, формування ефективної системи роботи з великими даними в біомедичних дослідженнях потребує узгодженого впровадження технологічних, методологічних та нормативних рішень. Одним з основних напрямів удосконалення є стандартизація форматів зберігання та обміну даними на основі міжнародних протоколів, зокрема HL7 FHIR, що забезпечують сумісність інформаційних систем і полегшують міжінституційну кооперацію. Важливо також підтримувати застосування відкритих API-інтерфейсів і гнучких структур метаданих, що сприяє підвищенню прозорості дослідницьких процесів і прискорює вторинне використання даних.

У сфері обробки даних рекомендується впроваджувати мультишарові аналітичні платформи, здатні інтегрувати омічні, клінічні, зображувальні й поведінкові дані в єдиний обчислювальний простір. Водночас доцільно посилити застосування explainable AI (пояснюваного штучного інтелекту), що дає змогу інтерпретувати результати автоматизованих висновків і зменшити

ризик помилкових або дискримінаційних рішень. Для забезпечення відтворюваності результатів необхідно підтримувати реєстрацію дослідницьких протоколів у відкритих системах, застосування контейнеризованих середовищ аналізу та систем контролю версій.

В аспекті регуляторної практики необхідно актуалізувати правові механізми захисту медичних даних, адаптуючи їх до викликів цифрової медицини. Це передбачає створення етичних комітетів для оперативної оцінки ризиків алгоритмізованих втручань, а також оновлення норм з інформованої згоди з урахуванням специфіки динамічного аналізу даних. Нарешті, доцільним є формування міждисциплінарних команд із залученням біоінформатиків, юристів, біоетиків і клініцистів для супроводу дослідницьких проєктів, що використовують великі дані. Такий підхід сприятиме не лише підвищенню об'єктивності й безпеки досліджень, а й зміцненню довіри до результатів, отриманих із застосуванням високотехнологічних інструментів у біомедичній науці.

Висновки. У ході дослідження встановлено, що великі дані є важливим чинником трансформації біомедичних досліджень, забезпечуючи перехід до індуктивного аналізу, адаптивного моделювання та персоналізованих клінічних рішень. Інтеграція омічних, реєстрових і поведінкових даних сприяє підвищенню точності доказової бази та прискорює її оновлення.

Виявлено основні проблеми, серед яких: технічна фрагментація даних, відсутність міжсистемної сумісності, етичні ризики при обробці персональної інформації та нормативна невизначеність щодо відповідальності за алгоритмізовані рішення. Обмеження також пов'язані з нерозвиненістю стандартів та недостатньою прозорістю аналітичних систем.

Рекомендовано впровадження єдиних стандартів (зокрема HL7 FHIR), розвиток explainable AI, створення багаторівневих аналітичних платформ та оновлення правових механізмів захисту медичних даних. Наголошено на необхідності міждисциплінарного супроводу досліджень.

Перспективи подальших розвідок охоплюють оцінювання якості медичних великих даних, розробку етичних протоколів для автоматизованих рішень і формування адаптивних цифрових екосистем, здатних забезпечити стабільну й безпечну інтеграцію аналітики в медичну практику.

Література:

1. Мойсеєнко М., Кузишин М., Туровська Л., Мазуренко Ю., Петришин М., Мазуренко О. Великі мовні моделі штучного інтелекту в медицині. *Сучасні інформаційні технології та інноваційні методики навчання в підготовці фахівців: методологія, теорія, досвід, проблеми*. 2024. Вип. 72, №. 72. С. 73–88. DOI: <https://doi.org/10.31652/2412-1142-2024-72-73-88> (дата звернення: 15.04.2025).
2. Cremin C., Dash S., Huang X. Big data: historic advances and emerging trends in biomedical research. *Current Research in Biotechnology*. 2022. Vol. 4, № 1. P. 138–151. DOI: <https://doi.org/10.1016/j.crbiot.2022.02.004> (date of access: 15.04.2025).
3. Yang X., Huang K., Yang D., Zhao W., Zhou X. Biomedical big data technologies, applications, and challenges for precision medicine: a review. *Global Challenges*. 2024. Vol. 8, № 1. Article 2300163. DOI: <https://doi.org/10.1002/gch2.202300163> (date of access: 15.04.2025).

4. Pei Y., Yang J. Biomedical applications of Big Data and Artificial Intelligence. *Bioengineering*. 2025. Vol. 12, № 2. P. 207. DOI: <https://doi.org/10.3390/bioengineering12020207> (date of access: 15.04.2025).
5. Taha K. Machine learning in biomedical and health big data: a comprehensive survey with empirical and experimental insights. *Journal of Big Data*. 2025. Vol. 12, № 1. Article 61. DOI: <https://doi.org/10.1186/s40537-025-01108-7> (date of access: 15.04.2025).
6. Rehman A., Naz S., Razzak I. Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities. *Multimedia Systems*. 2022. Vol. 28, № 4. P. 1339–1371. DOI: <https://doi.org/10.1007/s00530-020-00736-8> (date of access: 15.04.2025).
7. Sargiotis D. Leveraging AI and Big Data for Advancements in Biomedical Research. *Biomedical Journal of Scientific & Technical Research*. 2024. Vol. 57, № 4. P. 49586–49592. DOI: <https://doi.org/10.26717/BJSTR.2024.57.009048> (date of access: 15.04.2025).
8. Mirza B., Wang W., Wang J., Choi H., Chung N.C., Ping P. Machine learning and integrative analysis of biomedical big data. *Genes*. 2019. Vol. 10, № 2. Article 87. DOI: <https://doi.org/10.3390/genes10020087> (date of access: 15.04.2025).
9. Dash S., Shakyawar S.K., Sharma M., Kaushik S. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*. 2019. Vol. 6, № 1. P. 54. DOI: <https://doi.org/10.1186/s40537-019-0217-0> (date of access: 15.04.2025).
10. Banerjee A., Chakraborty C., Kumar A., Biswas D. Emerging trends in IoT and big data analytics for biomedical and healthcare technologies. *Handbook of Data Science Approaches for Biomedical Engineering*. 2020. Vol. 1, № 1. P. 121–152. DOI: <https://doi.org/10.1016/B978-0-12-818318-2.00005-2> (date of access: 15.04.2025).
11. Deng F., Huang J., Yuan X., Cheng C., Zhang L. Performance and efficiency of machine learning algorithms for analyzing rectangular biomedical data. *Laboratory Investigation*. 2021. Vol. 101, № 4. P. 430–441. DOI: <https://doi.org/10.1038/s41374-020-00525-x> (date of access: 15.04.2025).
12. Wang B., Li L. Research Progress in Biomedical Big Data. In: Ye, D.Q. (eds) *Progress in China Epidemiology*. Singapore: Springer, 2022. P. 325–344. DOI: https://doi.org/10.1007/978-981-19-2199-5_20 (date of access: 15.04.2025).
13. Ferretti A., Ienca M., Hurst S., Vayena E. Big data, biomedical research, and ethics review: new challenges for IRBs. *Ethics & Human Research*. 2020. Vol. 42, № 5. P. 17–28. DOI: <https://doi.org/10.1002/eahr.500065> (date of access: 15.04.2025).
14. Navale V., von Kaeppler D., McAuliffe M. An overview of biomedical platforms for managing research data. *Journal of Data, Information and Management*. 2021. Vol. 3, № 1. P. 21–27. DOI: <https://doi.org/10.1007/s42488-020-00040-0> (date of access: 15.04.2025).
15. Banerjee A., Chakraborty C., Kumar A., Biswas D. Emerging trends in IoT and big data analytics for biomedical and healthcare technologies. *Handbook of Data Science Approaches for Biomedical Engineering*. 2020. Vol. 1, № 1. P. 121–152. DOI: <https://doi.org/10.1016/B978-0-12-818318-2.00005-2> (date of access: 15.04.2025).
16. Apache Spark: website. 2025. URL: <https://spark.apache.org/> (date of access: 15.04.2025).
17. HL7 International: website. 2025. URL: <https://www.hl7.org/> (date of access: 15.04.2025).
18. FHIR – Fast Healthcare Interoperability Resources: website. 2025. URL: <https://www.hl7.org/fhir/> (date of access: 15.04.2025).
19. TensorFlow: website. 2025. URL: <https://www.tensorflow.org/> (date of access: 15.04.2025).
20. PACS – Picture Archiving and Communication System: Radiopaedia website. 2025. URL: <https://radiopaedia.org/articles/picture-archiving-and-communication-system-pacs> (date of access: 15.04.2025).
21. KNIME: website. 2025. URL: <https://www.knime.com/> (date of access: 15.04.2025).

22. Bioconductor: website. 2025. URL: <https://www.bioconductor.org/> (date of access: 15.04.2025).

23. What are Electronic Health Records (EHRs)?: HealthIT.gov: website. 2025. URL: <https://www.healthit.gov/topic/health-it-and-health-information-exchange-basics/what-are-electronic-health-records-ehrs> (date of access: 15.04.2025).

24. Clinical Decision Support (CDS): Agency for Healthcare Research and Quality: website. 2025. URL: <https://www.ahrq.gov/cpi/about/otherwebsites/clinical-decision-support/index.html> (date of access: 15.04.2025).

25. Artificial Intelligence in Cardiovascular Medicine: Mayo Clinic: website. 2025. URL: <https://www.mayoclinic.org/departments-centers/ai-cardiology/overview/ovc-20486648> (date of access: 15.04.2025).

26. Real-World Evidence: U.S. Food and Drug Administration: website. 2025. URL: <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence> (date of access: 15.04.2025).

27. Genomics and Medicine: National Human Genome Research Institute: website. 2025. URL: <https://www.genome.gov/health/Genomics-and-Medicine> (date of access: 15.04.2025).

References:

1. Moiseienko, M., Kuzyshyn, M., Turovska, L., Mazurenko, Y., Petryshyn, M., & Mazurenko, O. (2024). Velyki movni modeli shtuchnoho intelektu v medytsyni [Large language models of artificial intelligence in medicine]. *Suchasni informatsiini tekhnolohii ta innovatsiini metodyky navchannia v pidhotovtsi fakhivtsiv – Modern Information Technologies and Innovative Teaching Methods in Training Specialists*, 72(72), 73–88. Retrieved from: <https://doi.org/10.31652/2412-1142-2024-72-73-88>. [in Ukrainian]

2. Cremin, C., Dash, S., & Huang, X. (2022). Big data: Historic advances and emerging trends in biomedical research. *Current Research in Biotechnology*, 4(1), 138–151. DOI: <https://doi.org/10.1016/j.crbiot.2022.02.004>

3. Yang, X., Huang, K., Yang, D., Zhao, W., & Zhou, X. (2024). Biomedical big data technologies, applications, and challenges for precision medicine: A review. *Global Challenges*, 8(1), Article 2300163. DOI: <https://doi.org/10.1002/gch2.202300163>

4. Pei, Y., & Yang, J. (2025). Biomedical applications of big data and artificial intelligence. *Bioengineering*, 12(2), 207. DOI: <https://doi.org/10.3390/bioengineering12020207>

5. Taha, K. (2025). Machine learning in biomedical and health big data: A comprehensive survey with empirical and experimental insights. *Journal of Big Data*, 12(1), Article 61. DOI: <https://doi.org/10.1186/s40537-025-01108-7>

6. Rehman, A., Naz, S., & Razzak, I. (2022). Leveraging big data analytics in healthcare enhancement: Trends, challenges and opportunities. *Multimedia Systems*, 28(4), 1339–1371. DOI: <https://doi.org/10.1007/s00530-020-00736-8>

7. Sargiotis, D. (2024). Leveraging AI and big data for advancements in biomedical research. *Biomedical Journal of Scientific & Technical Research*, 57(4), 49586–49592. DOI: <https://doi.org/10.26717/BJSTR.2024.57.009048>

8. Mirza, B., Wang, W., Wang, J., Choi, H., Chung, N. C., & Ping, P. (2019). Machine learning and integrative analysis of biomedical big data. *Genes*, 10(2), Article 87. DOI: <https://doi.org/10.3390/genes10020087>

9. Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data*, 6(1), 54. DOI: <https://doi.org/10.1186/s40537-019-0217-0>

10. Banerjee, A., Chakraborty, C., Kumar, A., & Biswas, D. (2020). Emerging trends in IoT and big data analytics for biomedical and healthcare technologies. *Handbook of Data Science Approaches for Biomedical Engineering*, 1(1), 121–152. DOI: <https://doi.org/10.1016/B978-0-12-818318-2.00005-2>

11. Deng, F., Huang, J., Yuan, X., Cheng, C., & Zhang, L. (2021). Performance and efficiency of machine learning algorithms for analyzing rectangular biomedical data. *Laboratory Investigation*, 101(4), 430–441. DOI: <https://doi.org/10.1038/s41374-020-00525-x>
12. Wang, B., & Li, L. (2022). Research progress in biomedical big data. In D. Q. Ye (Ed.), *Progress in China Epidemiology* (pp. 325–344). Springer. DOI: https://doi.org/10.1007/978-981-19-2199-5_20
13. Ferretti, A., Ienca, M., Hurst, S., & Vayena, E. (2020). Big data, biomedical research, and ethics review: New challenges for IRBs. *Ethics & Human Research*, 42(5), 17–28. DOI: <https://doi.org/10.1002/eahr.500065>
14. Navale, V., von Kaeppler, D., & McAuliffe, M. (2021). An overview of biomedical platforms for managing research data. *Journal of Data, Information and Management*, 3(1), 21–27. DOI: <https://doi.org/10.1007/s42488-020-00040-0>
15. Banerjee, A., Chakraborty, C., Kumar, A., & Biswas, D. (2020). Emerging trends in IoT and big data analytics for biomedical and healthcare technologies. *Handbook of Data Science Approaches for Biomedical Engineering*, 1(1), 121–152. DOI: <https://doi.org/10.1016/B978-0-12-818318-2.00005-2>
16. Apache Software Foundation. (2025). Apache Spark. Retrieved from: <https://spark.apache.org/>
17. HL7 International. (2025). HL7 Standards. Retrieved from: <https://www.hl7.org/>
18. HL7 International. (2025). FHIR – Fast Healthcare Interoperability Resources. Retrieved from: <https://www.hl7.org/fhir/>
19. Google. (2025). TensorFlow. Retrieved from: <https://www.tensorflow.org/>
20. Radiopaedia.org. (2025). Picture Archiving and Communication System (PACS). Retrieved from: <https://radiopaedia.org/articles/picture-archiving-and-communication-system-pacs>
21. KNIME AG. (2025). KNIME Analytics Platform. Retrieved from: <https://www.knime.com/>
22. Bioconductor. (2025). Open software for bioinformatics. Retrieved from: <https://www.bioconductor.org/>
23. Office of the National Coordinator for Health Information Technology. (2025). What are Electronic Health Records (EHRs)?. Retrieved from: <https://www.healthit.gov/topic/health-it-and-health-information-exchange-basics/what-are-electronic-health-records-ehrs>
24. Agency for Healthcare Research and Quality. (2025). Clinical Decision Support (CDS). Retrieved from: <https://www.ahrq.gov/cpi/about/otherwebsites/clinical-decision-support/index.html>
25. Mayo Clinic. (2025). Artificial Intelligence in Cardiovascular Medicine. Retrieved from: <https://www.mayoclinic.org/departments-centers/ai-cardiology/overview/ovc-20486648>
26. U.S. Food and Drug Administration. (2025). Real-World Evidence. Retrieved from: <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>
27. National Human Genome Research Institute. (2025). Genomics and Medicine. Retrieved from: <https://www.genome.gov/health/Genomics-and-Medicine>