УДК 617.713-002-071:004.8

Можливості штучного інтелекту у діагностиці кератоконусу з використанням структурованого алгоритму

Безкоровайна І. М. [1], д-р мед. наук, професор; **Гонтар А. Р.** [1], лікар-інтерн;
Наконечний Д. О. [2], лікар-офтальмолог; **Іванченко А. Ю.** [1], PhD, асистент

[1] Полтавський державний медичний університет
Полтава (Україна)

[2] Приватний офтальмологічний центр «Світогляд»
Полтава (Україна)

Мета. *Оцінка ефективності застосування штучного інтелекту у діагностиці кератоконусу на основі кератотопографічних карт, з подальшим визначенням стадії захворювання та дослідження впливу рівня стандартизації клінічного алгоритму на точність і стабільність результатів.*

Матеріал та методи. *Проведено ретроспективний аналіз 59 кератотопографічних карт із бази даних ACE DIAGNOSTIC PLATFORM фірми Baush+Lomb, 2021 року випуску, отриманих від 37 пацієнтів. Дослідження здійснювалося у три етапи, які моделювали різні рівні інтеграції ШІ – від повної автономії до роботи за жорстко структурованим алгоритмом діагностики, розробленим авторами (авторське право на твір № 139012 від 26.08.2025) [12]. Для встановлення діагнозу та стадії використовувалась адаптована мовна модель ChatGPT (OpenAI, 2023). Статистична обробка результатів дослідження проводилася за допомогою програмного забезпечення Microsoft Excel 2019 із використанням вбудованих статистичних функцій і надбудови Analysis ToolPak з використанням моделі бінарної та багатокласової класифікації, порівнювалися точність, чутливість, специфічність, прецизійність та F1-міра на кожному етапі.*

Результати. *Встановлено, що точність моделі на першому етапі склала 79,7%, однак лише 33,3% випадків були правильно стадійовані. На другому етапі загальна точність зросла до 93,2%, а точність визначення стадії – до 74,36%. Найвищі показники отримано на третьому етапі: чутливість – 100%, специфічність – 95%, точність – 98,0%, точність стадіювання – 89,7%.*

Висновки. *Штучний інтелект продемонстрував потенціал для автоматизованої діагностики кератоконусу, однак ефективність реалізується повною мірою лише за умови чіткого структурованого алгоритму. Самостійне рішення ШІ без вказівок призводить до високої варіативності результатів. Впровадження алгоритмізованої логіки дозволяє значно підвищити стабільність і точність системи.*

Ключові слова:
рогівка, штучний інтелект, ChatGPT, кератоконус, кератотопографія

Вступ. Кератоконус – це прогресуюче дегенеративне захворювання рогівки, що веде до зниження гостроти зору та можливої інвалідизації. Це захворювання належить до категорії так званих «тихих патологій», оскільки на початкових етапах може тривалий час залишатися не діагностованим або сплутаним із простими формами міопії чи астигматизму [1-4]. Прогресування кератоконусу вимагає своєчасного виявлення, адже з плином часу захворювання переходить у більш тяжкі стадії, які нерідко потребують хірургічного втручання, зокрема крос-лінкінгу, імплантації сегментів чи навіть кератопластики [5-6].

Одним із головних методів діагностики кератоконусу є комп'ютерна кератотопографія – метод візуалізації рельєфу рогівки, який дає змогу оцінити кривизну, товщину, елевацію передньої і задньої поверхонь, а також симетричність рогівкових структур [4,6]. Однак попри об'єктивність отриманих цифрових зображень, їх інтерпретація залишається суб'єктивною, що залежить від рівня кваліфікації, досвіду лікаря, наявності клінічного мислення та алгоритмічного підходу. У багатьох клініках, особливо на первинній ланці, ці чинники відсутні або недостатньо реалізовані, що призводить до запізнілої діагностики кератоконусу, чи надмірної настороженості.

Останні роки ознаменувалися бурхливим розвитком штучного інтелекту (ШІ) у медицині. У сфері офтальмології особливої уваги набули дослідження, що спрямовані на виявлення діабетичної ретинопатії, глаукоми, вікової макулодистрофії, а також кератоконусу за допомогою глибокого навчання та нейронних мереж [7-10]. Однак жодне з них не набуло великого поширення.

Попередні дослідження [8-11] показали, що великі мовні моделі та нейромережеві архітектури мають

високу здатність до аналізу кератотопографічних зображень, однак більшість із них поки не впроваджені в клінічну практику через складність стандартизації рішень, брак прозорості логіки та ризик помилкових висновків. У цьому контексті надзвичайно важливим є поєднання ШІ з чітким алгоритмом логічного аналізу, який враховує всі клінічно релевантні параметри та дозволяє уникнути надмірної автономії системи.

У зв'язку з цим, розробка структурованого підходу до автоматизованої діагностики кератоконусу із залученням штучного інтелекту стає надзвичайно актуальною. Визначення як самого факту наявності захворювання, так і його стадії має ключове значення для прогнозу та вибору тактики лікування. Саме це завдання і було покладене в основу даного дослідження.

Мета дослідження. Оцінка ефективності застосування штучного інтелекту у діагностиці кератоконусу на основі кератотопографічних карт, з подальшим визначенням стадії захворювання та дослідження впливу рівня стандартизації клінічного алгоритму на точність і стабільність результатів.

Матеріал і методи

Дослідження носило ретроспективний характер і проводилося на кафедрі оториноларингології з офтальмологією Полтавського державного медичного університету з використанням бази кератотопографічних даних, зібраних за допомогою пристрою ACE DIAGNOSTIC PLATFORM 2021 року випуску фірми BAUSCH + LOMB у приватній клініці «Світогляд» (м. Полтава). Усього було проаналізовано 59 кератотопографічних карт, що відповідали якісним критеріям для цифрової обробки: чіткість зображення, коректність центрування та відсутність артефактів, що могли б спотворити результати.

До дослідження було включено 37 пацієнтів (59 очей). З них 27 чоловіків та 10 жінок віком від 20 до 54 років. Основна група становила 39 очей пацієнтів із клінічно підтвердженим діагнозом кератоконусу (I–IV стадія за результатами лікарського консиліуму), тоді як контрольна група включала 20 очей здорових осіб без ознак патології рогівки. Виключення складали пацієнти з наявністю пеллюцидної маргінальної дегенерації, рубцевими змінами після запальних захворювань, травм та з післяопераційними змінами рогівки. Для аналізу використовувалася мовна модель ChatGPT (OpenAI, 2023) із доступом до інтерфейсу, що дозволяв подавати кератотопографічні карти у вигляді зображень безпосередньо для автоматичного розпізнавання та інтерпретації. Аналіз відбувався у три етапи, кожен з яких імітував різний рівень автономності ШІ у клінічній діагностиці.

На першому етапі ChatGPT отримував лише візуальні або числові дані з кератотопографічної карти (PachyMin, найтонша точка рогівки) – зміщення найтоншої точки по осі Y, симетрія Anterior Axial Curvature, елевація передньої поверхні, елевація задньої поверхні, Q-value (показник асферичності рогівки), Kmax (показник максимальної кривизни рогівки, астигматизм) – без жодного коментаря або вказівок щодо норм, логіки діагностики, чи структурованого алгоритму. Метою було визначити, наскільки мовна модель здатна самостійно класифікувати кератоконус та його стадії, використовуючи лише власну базу знань і логічні припущення.

На другому етапі моделі надавалися орієнтовні вказівки щодо основних діагностичних показників та діапазонів ризику, без прямого формулювання діагностичного алгоритму. Наприклад, модель знала, що PachyMin< 470 мкм є потенційним маркером кератоконусу, або що Kmax> 47,2 D може свідчити про патологію. Однак інтерпретація цих даних і формування остаточного діагнозу залишалася за ШІ.

На третьому етапі модель діяла суворо відповідно до розробленого нами структурованого трирівневого алгоритму (рівень 1 – окремий аналіз показників, рівень 2 – визначення патології, рівень 3 – стадіювання) (авторське право на твір № 139012 від 26.08.2025) [12], що передбачав поетапний аналіз кожного параметра з урахуванням чітких порогових значень і правил переходу між рівнями.

Дані обчислювалися на основі порівняння результатів ChatGPT з консиліумним клінічним діагнозом. Всі розрахунки проводилися вручну з подальшим перехресним контролем.

Аналіз даних проводився шляхом розв'язання завдань бінарної та багатокласової класифікації. Якість класифікації оцінювали за допомогою набору метрик: для бінарної класифікації використовували точність, чутливість, специфічність, прецизійність та F1-міру; для багатокласової класифікації використовували макросередні значення прецизійності, чутливості та F1-міри, а також загальну точність. Дослідження проводилося в три етапи: навчання на нерозмічених даних; навчання на розмічених даних без стадіювання; навчання на розмічених даних зі стадіюванням. Порівнювали класифікаційні метрики, отримані на всіх етапах. Стабільність рішення оцінювали шляхом аналізу показника F1-міри на кожному етапі роботи моделі. Для математичної обробки даних застосовано Microsoft Excel 2019 з використанням вбудованих статистичних функцій і надбудови Analysis ToolPak. У процесі аналізу враховувалися кількісні показники істинно позитивних, хибно позитивних, істинно негативних та хибно негативних результатів, на основі яких обчислювалися відповідні метрики діагностичної ефективності. Для кожного параметра додатково визначалися 95%-ні довірчі інтервали (ДІ) з метою оцінки статистичної надійності отриманих результатів.

Результати

На першому етапі, коли мовна модель діяла без будь-яких підказок, загальна точність становила 79,7%. Із тридцяти дев’яти випадків кератоконусу

ШІ правильно діагностував тридцять один (чутливість – 79,5%; 95% ДІ: 64,5–89,2), тоді як із двадцяти контрольних зображень шістнадцять були коректно віднесені до норми (специфічність – 80,0%; 95% ДІ: 58,4–91,9). Однак найслабшою ланкою залишилося стадіювання: лише тринадцять випадків із тридцяти дев'яти були точно співвіднесені з клінічною стадією (33,3%), причому помилки мали доволі хаотичний характер. В таблиці представлені результати обробки кератотопограм ШІ без надання йому будь-яких інструкцій (1-й етап). При повторному аналізі тих самих карт ChatGPT виявив високу мінливість: у сімнадцяти зразках модель змінювала попередній висновок, що свідчило про відсутність стабільності рішень.

Після введення орієнтовних меж діагностично значущих показників (2-й етап) відбулося суттєве покращення усіх метрик. Модель коректно ідентифікувала 36 патологічних карт (чутливість – 92,3%; 95% ДІ: 79,7–97,4) та 19 із 20 здорових (специфічність – 95,0%; 95% ДІ: 76,4–99,1). Загальна точність зросла до 93,2%, а F1-міра до 94,7%, що продемонструвало ефективність навіть часткового стандартизування діагностичних критеріїв. Одночасно точність визначення стадії збільшилася майже втричі, до 74,4% (32 коректних стадіювання із 39). Результати дослідження на другому етапі представлено у таблиці 1. Проте стабільності рішень не було досягнуто: при завантаженні для повторного дослідження дев'ять карт отримали інший ступінь стадіювання.

На цьому етапі проведення дослідження виникли технічні труднощі, пов'язані з обмеженнями безкоштовної версії мовної моделі ChatGPT. Зокрема, обмежена кількість запитів, обсяг контексту для аналізу та неможливість обробки великої кількості зображень в межах однієї сесії унеможливлювали повноцінне тестування. Перехід на розширену версію ChatGPTPlus (платна підписка від OpenAI), що відкриває доступ до потужнішої моделі GPT-4, дозволив продовжити дослідження, зняти технічні обмеження та забезпечити повноцінну реалізацію другого й третього етапів аналізу.

На третьому етапі дослідження, де ChatGPT функціонував суворо в межах розробленого нами трирівневого алгоритму, модель безпомилково розпізнала всі 39 випадків кератоконусу (чутливість – 100,0%; 95% ДІ: 91,0–100,0) і лише в одному випадку помилково віднесла здорове око до патології, що дало специфічність 95,0%. У випадку здорового ока, який було визначене алгоритмом як субклінічний, числові параметри кератотопограми становили: PachyMin – 580 мкм, зміщення найтоншої точки по осі (y) –0,56, симетрія на карті Anterion Axial Curvature у межах норми, Q-value= –0,24, Kmax – 43,64D, астигматизм – 0,54D. За клінічною оцінкою та топографічними картами це око відповідало нормі. Загальна точність досягла 98,0% (95% ДІ: 91,0–99,7), а F1-міра – 98,7%. Надзвичайно показовим виявилося стадіювання: у 35 із 39 карт (89,7%; 95% ДІ: 76,4–95,9) номер стадії повністю збігся з консиліумним клінічним висновком, що представлено в таблиці.

Додатково оцінювалася стабільність рішення. При повторному обробленні всього масиву карт протягом тижня ChatGPT, працюючи згідно з алгоритмом третього етапу, зберігав ідентичні результати у 100% випадків, що підтвердило відсутність розбіжностей, пов'язаних із стохастичною природою мовної моделі.

Обговорення

Згідно з результатами численних досліджень, штучний інтелект уже демонструє високу ефективність у діагностичних завданнях офтальмології, часто не поступаючись лікарям, проте його широке клінічне впровадження потребує подальших досліджень, зокрема щодо узагальнюваності, інтерпретованості та адаптації до реальних умов [8-10, 13]. Результати нашого дослідження продемонстрували значний потенціал мовних моделей, зокрема ChatGPT, у сфері аналізу та інтерпретації кератотопографічних даних. Важливо враховувати, що сам по собі штучний інтелект не забезпечував діагностичну безпомилковість. Лише при наявності чітких логічних рамок модель може виконувати функції, наближені до клінічних рішень лікаря. Це узгоджується з результатами закордонних досліджень, які відзначають ефективність моделей на базі глибокого навчання лише при обмеженому і контрольованому впливі людського фактора або навпаки – при повній стандартизації вхідних даних і логіки [2, 4, 6-8].

Цікаво, що автономна робота ШІ на першому етапі демонструвала прийнятні показники чутливості та специфічності, однак виявила слабке місце – стадіювання. Це пояснюється тим, що без алгоритмічної жорсткості ChatGPT схильний до надінтерпретації або пропуску важливих параметрів, особливо у сумнівних

Таблиця 1. Результати діагностичної точності моделі штучного інтелекту на трьох етапах дослідження.

	Істинно позитивні (TP)	Істинно негативні (TN)	Хибно позитивні (FP)	Хибно негативні (FN)	Чутливість	Специфічність	Загальна точність	Прецизійність	F1-міра	Точність визначення стадії
1-й етап	31	16	4	8	79,5%	80,0%	79,7%	88,6%	83,8%	33,33%
2-й етап	36	19	1	3	92,3%	95,0%	93,2%	97,3%	94,7%	74,36%
3-й етап	39	19	1	0	100,0%	95,0%	98,0%	97,5%	98,7%	89,74%

випадках. Саме тому впровадження структури на рівні діагностичних рішень не лише бажане, а й необхідне. Водночас третій етап, що базувався на розробленому трирівневому алгоритмі, дозволив моделі чітко визначати ключові комбінації ризиків і відокремлювати субклінічні, граничні та клінічно виражені форми кератоконусу.

Важливим практичним аспектом стала стабільність повторного аналізу. В умовах клінічної реальності, особливо при динамічному моніторингу пацієнтів, стабільність інтерпретації має критичне значення. Нестабільність, виявлена на перших етапах, підтверджує, що ШІ в «творчому» режимі не може бути єдиною підставою для постановки діагнозу. Лише після введення фіксованих правил алгоритм забезпечив повну відтворюваність результатів, що свідчить про зниження ймовірності випадкових помилок і підвищення клінічної надійності.

Варто також зазначити, що не всі карти були однозначними: частина пацієнтів мала показники, які не дозволяли навіть лікарю виставити чіткий діагноз без подальшого спостереження. Саме тому в алгоритм було закладено категорію «пацієнт під увагою», що відповідає концепції ранньої стратифікації ризиків. Це дозволяє мінімізувати випадки хибнопозитивних рішень, які могли б призвести до передчасного лікування або психологічного тиску на пацієнта.

Враховуючи зростаючу роль штучного інтелекту в сучасній медицині, результати цього дослідження та створений алгоритм можуть слугувати внеском у подальші напрацювання у сфері гібридних експертних систем, де взаємодія між клініцистом і ШІ ґрунтуватиметься на принципах доповнення та взаємного контролю, з незмінним збереженням ключової ролі лікаря у процесі прийняття рішень.

На завершення потрібно сказати, що важливо впроваджувати штучний інтелект у діагностику кератоконусу, так як це підвищує об'єктивність медичних рішень, мінімізує вплив людського фактора та скорочує час постановки діагнозу. Але робити це необхідно з використанням правильно сформованого алгоритмізованого підходу, що дозволяє значно підвищити точність стадіювання та стабільність діагностичних рішень і підвищує клінічну надійність ШІ у практичній офтальмології. Перспективи подальших досліджень включають розширення бази даних, інтеграцію з іншими методами обстеження рогівки, а також розробку гібридних експертних систем, де клінічне рішення поєднується з алгоритмізованою аналітикою ШІ.

Література

1. **Jacinto S, Carracedo G, Suzaki A, Villa-Collar C, J. Vincent S, Wolffsohn SJ.** Keratoconus: An updated review. Science Direct. 2022; 45(3):101559. doi: 10.1016/j.clae.2021.101559
2. **Lin SR, Ladas JG, Bahadur GG, Al-Hashimi S, Pineda R.** A Review of Machine Learning Techniques for Keratoconus Detection and Refractive Surgery Screening. Seminars in Ophthalmology. 2019;34(4):317–326. doi: 10.1080/08820538.2019.162081
3. **Ferdi AC, Kandel H, Nguyen V, Tan J, Arnalich-Montiel F, Abbondanza M, Watson SL.** Five-year corneal cross-linking outcomes: A Save Sight Keratoconus Registry Study. Clin Exp Ophthalmol. 2023 Jan;51(1):9-18. doi: 10.1111/ceo.14177.
4. **Bui AD, Truong A, Pasricha ND, Indaram M.** Keratoconus Diagnosis and Treatment: Recent Advances and Future Directions. Clinical Ophthalmology. 2023;16(17):2705-2718. doi: 10.2147/OPTH.S 392665
5. **Niazi S, Gatzioufas Z, Doroodgar F, Findl O, Baradaran-Rafii A, Liechty J, et al.** Keratoconus: exploring fundamentals and future perspectives – a comprehensive systematic review. Sage Journals. 2024;16 16:25158414241232258.
6. **Li Z, Wang L, Wu X, Jiang J, Qiang W, Xie H, et al**. Artificial intelligence in ophthalmology: The path to the real-world clinic. Cell Reports Medicine. 2023;4(7):101095. doi: 10.1016/j.xcrm.2023.101095
7. **Azzahra Afifah, Fara Syafira, Putri Mahirah Afladhanti, Dini Dharmawidiarini.** Artificial intelligence as diagnostic modality for keratoconus: A systematic review and meta-analysis. Journal of Taibah University Medical Sciences. 2023;19(2): 296-303.
8. **Wan Q, Wei R, Ma K, Yin H, Deng YP, Tang J.** Deep Learning-Based Automatic Diagnosis of Keratoconus with Corneal Endothelium Image. Ophthalmol Ther. 2023 Dec;12(6):3047-3065. doi: 10.3390/diagnostics 13162715.
9. **Muhsin ZJ, Qahwaji R, Al Shawabkeh M. et al.** Smart decision support system for keratoconus severity staging using corneal curvature and thinnest pachymetry indices. Eye and Vis. 2024;11(1):28. doi: 10.1186/s 40662-024-00394-1.
10. **Yousefi S, Yousefi E, Takahashi H, Hayashi T, Tampo H, Inoda S, Arai Y, Asbell P.** Keratoconus severity identification using unsupervised machine learning. PLo S One. 2018;13(11) e 0205998. doi: 10.1371/journal.pone.0205998.
11. **Zhamardiy VO, Shkola OM, Okhrimenko IM, Strelchenko OG, Aloshyna AI. et al**. Checking of the methodical system efficiency of fitness technologies application in students' physical education. Wiad Lek. 2020;73(2):332–341. doi: 10.36740/WLek202002125
12. https://surl.li/ztvper
13. **Nevska A. et al.** Assessing the possibility of using portable and stationary non-mydriatic fundus cameras for diabetic retinopathy screening assisted by an artificial intelligence-based software platform in primary care. Journal of Ophthalmology (Ukraine). 2025; 6: 22–26. DOI:https://doi.org/10.31288/oftalmolzh202462226.

Відомості про авторів та розкриття інформації

***Внесок кожного автора в роботу**. Безкоровайна І.М. - дизайн, методологія, написання, рецензування та редагування; Гонтар А.Р. - збір данних, написання алгоритму, написання статті; Наконечний Д.О. - збір данних, аналіз застосування алгоритму; Іванченко А.Ю. - редагування, статистичний аналіз. Усі автори проаналізували результати та погодили кінцевий варіант рукопису.*

Відмови від відповідальності: *висловлені у поданій статті думки є власними поглядами авторів, а не офіційними позиціями установи.*

Джерела підтримки*. Зовнішні джерела фінансування відсутні.*

Конфлікт інтересів. *Автори свідчать про відсутність конфліктів інтересів, які б могли вплинути на їх думку стосовно предмету чи матеріалів, описаних та обговорених в даному рукописі.*

Заява про доступність даних. *Всі дані, отримані або проаналізовані під час цього дослідження, включені в цю опубліковану статтю.*

Список скорочень: *ШІ – штучний інтелект, ChatGPT – Generative Pre trained Transformer, AI – artificial intelligence, ДІ – довірчі інтервали.*

Надійшла 28.07.2025