

УДК (UDC) 004.89+519.6

**Тюрдьо  
Іван Миколайович**

*аспірант факультету математики і інформатики, кафедра  
прикладної математики  
Харківський національний університет імені В. Н. Каразіна, майдан  
Свободи, 4, м. Харків, 61022  
e-mail: [ivan.turdio@karazin.ua](mailto:ivan.turdio@karazin.ua)  
<https://orcid.org/0009-0001-7315-3628>*

**Седюк  
Анастасія Денисівна**

*студентка факультету математики і інформатики, кафедра  
прикладної математики  
Харківський національний університет імені В. Н. Каразіна, майдан  
Свободи, 4, м. Харків, 61022  
e-mail: [seduknasta7@gmail.com](mailto:seduknasta7@gmail.com)  
<https://orcid.org/0009-0008-0026-693X>*

**Кізілова  
Наталія Миколаївна**

*доктор фізико-математичних наук, професор кафедри прикладної  
математики  
Харківський національний університет імені В. Н. Каразіна, майдан  
Свободи, 4, м. Харків, 61022  
e-mail: [kizilova@karazin.ua](mailto:kizilova@karazin.ua)  
<https://orcid.org/0000-0001-9981-7616>*

## Математичне моделювання динаміки зростання пухлини для вибору персоналізованої терапії

**Мета роботи:** виконати аналіз сучасних підходів до математичного моделювання росту пухлин та прогнозування їхньої динаміки із застосуванням класичних детермінованих моделей і методів машинного навчання, а також визначити перспективи їх використання у сучасній математичній онкології та персоналізованій протипухлинній терапії.

**Методи дослідження:** аналіз і систематизація сучасних наукових публікацій з математичної онкології; використання методів математичного моделювання росту пухлин (експоненціальні, логістичні, моделі Гомпертца та Берталанфі); статистичний аналіз клінічних даних; застосування методів машинного навчання для регресійного аналізу та прогнозування динаміки росту пухлин на основі поздовжніх МРТ-даних відкритого набору LUMIERE.

**В результаті** дослідження виконано огляд і порівняльний аналіз класичних математичних моделей росту пухлин та їхніх модифікацій, що використовуються для опису біологічних процесів проліферації та обмеження росту пухлинної тканини. Проведено попередню обробку та аналіз клінічних і візуалізаційних даних, що включають об'єми різних компонентів пухлини. Здійснено моделювання індивідуальних траєкторій росту пухлин із використанням регресійних моделей та ансамблевих методів машинного навчання, зокрема Random Forest. Показано, що методи машинного навчання забезпечують більш стійке та точне прогнозування складної динаміки росту пухлин порівняно з класичними моделями у випадку високої варіабельності даних.

**Висновки:** поєднання класичних математичних моделей росту пухлин із сучасними методами машинного навчання є перспективним напрямком розвитку математичної онкології. Такий підхід дозволяє підвищити точність прогнозування індивідуальної динаміки пухлин та створює основу для розробки персоналізованих стратегій лікування. Отримані результати свідчать про доцільність подальшого використання гібридних моделей у дослідженнях з прецизійної медицини та персоналізованої протипухлинної терапії.

**Ключові слова:** математична онкологія, моделі росту пухлин, модель Гомпертца, модель Берталанфі, машинне навчання, Random Forest, прогнозування, персоналізована терапія, прецизійна медицина, клінічні дані.

**Як цитувати:** Тюрдьо І.М., Седюк А.Д., Кізілова Н. М. Математичне моделювання динаміки зростання пухлини для вибору персоналізованої терапії. *Вісник Харківського національного університету імені В. Н. Каразіна, серія Математичне моделювання. Інформаційні технології. Автоматизовані системи управління.* 2026. вип. 69. С.82-100.

<https://doi.org/10.26565/2304-6201-2026-69-07>

**How to quote:** I. Tiurdo, A. Sediuk, N. Kizilova “Mathematical modeling of tumor growth dynamics for personalized therapy selection”, *Bulletin of V. N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 69, pp. 82-100, 2026. <https://doi.org/10.26565/2304-6201-2026-69-07> [in Ukrainian]

## Вступ

Онкологічні захворювання є однією з провідних причин смертності у світі. За даними ВООЗ [1], загальне число нових випадків раку у 2019 р. було навколо 18 млн. У 2022 р. ця цифра зросла до 19,3 млн, а прогнозування на 2025 та подальші роки складає більше 19 млн. щорічно [2]. В Україні проживає більше 1,3 млн. хворих на рак, і щорічно додається приблизно 160 тисяч нових випадків хвороби. Сучасні методи діагностики і лікування інтенсивно розроблюються у фармакології, медицині і біоінженерних науках [3, 4, 5]. Також математичне моделювання, яке є високоефективним інструментом дослідження різних біологічних процесів, дозволяє прогнозувати поведінку складних систем на основі дослідних даних і розробляти інноваційні підходи для діагностики та лікування захворювань.

Однією з ключових задач математичного моделювання є опис зростання пухлин за допомогою низки параметрів, які характеризують фізіологічні особливості ракових клітин, включаючи їх розмір, форму, симетрію, текстуру тощо. Цінність таких моделей полягає в їх здатності враховувати як локальні зміни в окремих клітинах, так і глобальну поведінку пухлини в організмі, у тому числі її відгук на хіміо- або радіотерапію. Сучасна медицина все більше покладається на математичне моделювання як на спосіб інтерпретації складних біологічних процесів. В умовах бурхливого розвитку машинного навчання й аналітики даних, класичні математичні моделі можуть поєднуватись із алгоритмами штучного інтелекту для створення більш точних і адаптивних систем прогнозування.

Актуальність обраної теми полягає у зростаючій потребі в математичних інструментах, які дозволяють точно описувати, аналізувати і прогнозувати динаміку зростання пухлини під впливом обраної терапії.

Метою роботи є розробка і тестування сучасних математичних алгоритмів для аналізу реальних клінічних даних зростання пухлин для удосконалення математичного моделювання динаміки індивідуального розвинення пухлини до та під час хіміотерапії.

## 1. Огляд літератури

Зростання пухлини – це складне явище, яке включає низку взаємопов'язаних фізіологічних процесів, починаючи з появи ракових клітин, їх проліферації (швидке зростання чисельності за рахунок клітинного ділення) та організації у просторі як пухлини, з розвинутою системою кровопостачання. Відповідні математичні моделі поділяються на найпростіші, які описують зміну з часом маси (або об'єму, або чисельності клітин), та більш складні, які беруть до уваги неоднорідність популяції клітин, відсутність або наявність метаболічних обмежень (недостатність кисню та живильних речовин), конвекційний (з рухом крові, лімфи, тканинної рідини) та/або дифузійний транспорт хімічних речовин, та інші аспекти [6].

Зі зростанням пухлини її клітини починають відчувати дефіцит кисню та поживних речовин, що стимулює ангиогенез — формування нових судин під дією фактору зростання [7, 8, 9]. Швидка агресивна проліферація призводить до того, що клітини всередині пухлини не отримують достатнє живлення і відмирають, утворюючи некротичне ядро пухлини, яке може призвести до її руйнування. Але паралельно розвивається здатність пухлинних клітин до інвазії та метастазування: вони руйнують міжклітинні зв'язки, проникають у сусідні тканини й судини, що значною мірою зумовлює агресивність перебігу хвороби [7]. Таким чином, якщо рухливість пухлинних клітин збільшується зі зростанням їх щільності, це дає можливість пухлини рухатися у різних напрямках а також пересуватися по організму з кровотоком, що дає можливість пухлині розвиватися далі, незалежно від її початкового розміру. Якщо рухливість клітин зменшується зі щільністю, будь-яка популяція ракових клітин нижче деякого порогу розміру зникне з часом (ефект Алле [10]). Крім того, зростання пухлини супроводжується формуванням складного мікрооточення: імунні клітини, фібробласти, ендотелій та сигнальні молекули взаємодіють із пухлинними клітинами, змінюючи перебіг процесу [7].

У багатьох випадках траєкторія зростання набуває S-подібної форми: після експоненціального старту темпи зростання сповільнюються і формується плато внаслідок дефіциту ресурсів або

терапевтичного впливу [11, 12]. Для опису зростання були запропоновані різні математичні моделі, які дають у якості розв'язку відповідного рівняння саме S-подібні криві.

Перші математичні моделі зростання ракових пухлин з'явилися ще 200 років тому. У 1825 р. англійський математик Бенджамін Гомпертц запропонував балансове рівняння для опису смертності населення, яке пізніше було використано для опису кривих зростання пухлин і їх екстраполяції на одну ракову клітину [13]. У 1838р. бельгійський математик П'єр-Франсуа Ферхюльст запропонував балансову рівняння для опису динаміки зростання населення, яке теж пізніше почало використовуватися для опису зростання пухлин. Значно пізніше, у 1938 р. австрійський біолог Людвіг фон Бергаланфі запропонував балансове рівняння для опису індивідуального зростання організму (маси або розміру, довжини тварини), яке теж пізніше знайшло використання у дослідженнях динаміки пухлин [14].

У 1954 р. була запропонована модель багатоетапного розвинення специфічних пухлин за рахунок накопичення ймовірностей мутацій і переродження, яка спиралася на статистичні дані розподілу смертності від різних типів раку за віком хворих [15]. У 1966 р. Артур Буртон запропонував першу дифузійну модель зростання пухлини [16], яка була пізніше розвинута у роботах відомого математика Харві Грінспена [17]. Одночасно розвивався підхід, який детальніше розглядав перенос живильних речовин і кисню по судинах та міжклітинному простору до здорових і ракових клітин. У 1937 р. були опубліковані дві роботи Теорелла про розповсюдженням ліків по системах і тканинах організму [18,19], на основі яких у 1961 р. американський математик Ричард Белман запропонував компартментальну модель хіміотерапії [20].

Таким чином, наприкінці 1960-х – на початку 1970-х років склалася нова міждисциплінарна галузь математична онкологія - наука, яка на основі статистичних даних і відомих біологічних закономірностей динаміки пухлин будує математичні моделі, які дозволяють глибше зрозуміти процес утворення і зростання пухлин від молекулярного до клітинного і тканинного рівня [21, 22, 23]. Генетичні й епігенетичні порушення призводять до активації онкогенів, пригнічення генів-супресорів, втрати контролю над апоптозом і появи гетерогенності клітин [7]. Така гетерогенність забезпечує пухлині здатність до адаптації та стійкість до лікування.

Математичне моделювання ґрунтується на звичайних диференціальних рівняннях (ЗДР) та рівняннях у часткових похідних, які описують не тільки збільшення або зменшення маси пухлини, але й динаміку утворення нових кровоносних судин (ангіогенез), вплив механічного стискання, розповсюдження метастаз та відгук пухлини та різні види протипухлинної терапії. Моделі розвитку гетерогенних пухлин розглядають два і більше ЗДР для популяцій клітин різних типів [24], наприклад, резистивних до ліків та сприйнятливих ракових клітин; проліферуючих та «сплячих» клітин [25]; проліферуючих та відмираючих (апоптоз) клітин [26, 27]; або клітин пухлини та природних антипухлинних клітин (лейкоцитів певних типів) [28]. Цікаві моделі для гетерогенних пухлин виникають у разі наявності запізнення у часі між поведінкою різних популяцій [29, 26], а також проблеми стійкості зростання [30]. Крім того, існують моделі інших типів: дискретні [31], моделі клітинних автоматів [32].

Сучасна математична онкологія активно розвивається і використовує як статистичні методи аналізу і аналітики «великих даних», так і обчислювальні методи (скінченних елементів, скінченних об'ємів та ін.), які враховують вплив мікрооточення пухлини, у тому числі механічне стискання [33], а також використовують підходи штучної інтелігенції. Однак, найпростіші балансові математичні моделі у вигляді ЗДР для одного або кількох параметрів залишаються актуальними. Це пов'язано з тим, що, як у лабораторних дослідженнях *in vivo*, так і у спостереженнях за пацієнтами, використовують 2d або 3d зображення пухлини, на яких можна вимірювати динаміку змін її розмірів (висота  $\times$  ширина  $\times$  довжина), за якими можна обчислити об'єм пухлини та у подальшому використовувати ЗДР для об'єму як еліпсоїду або для маси пухлини, де  $\rho$  - середня густина,  $t$  – час, який вимірюються у добах (тижнях, місяцях) [34].

Вагому роль у сучасній онкології відіграє кількісний аналіз пухлинних параметрів. Завдяки розвитку медичної візуалізації (КТ, МРТ) і автоматизованої радіоміки стало можливим неінвазивно оцінювати об'єм, структуру, текстуру та функціональні властивості пухлин [35]. Особливо це актуально для гліобластоми, де сучасні протоколи МРТ із автоматизованим аналізом дозволяють будувати персоналізовані прогностичні моделі та оцінювати ефективність лікування у динаміці [35]. Усе це формує фундамент для побудови сучасних математичних моделей, що кількісно описують динаміку пухлини й забезпечують основу для персоналізованої медицини.

## 2. Класифікація моделей

Математичні моделі росту пухлин є потужним інструментом для кількісного опису динаміки розвитку злоякісних новоутворень. Вони дозволяють досліджувати механізми проліферації клітин, оцінювати ефективність лікування та прогнозувати перебіг хвороби. У літературі описано багато підходів, від простих детермінованих моделей на основі диференціальних рівнянь до складних стохастичних і мультіагентних систем. У цьому розділі розглянуто класичні детерміновані моделі росту пухлини, а також обґрунтовано вибір регресійного підходу в контексті роботи з клінічними даними.

### 2.1 Лінійні моделі

Лінійні моделі є найпростішими і відповідають зростанню пухлини з постійною швидкістю

$$\frac{dV}{dt} = b, \quad V(t_0) = V_0, \quad (1)$$

де  $V(t)$  - об'єм (маса) пухлини,  $b$  – швидкість зростання,  $V_0$  - значення об'єму у заданий початковий момент спостережень.

Вибір об'єму відповідає спостереженням (*in vivo*), коли є можливість отримувати тільки дані візуалізації пухлини в організмі за допомогою КТ, МРТ або УЗ досліджень. Розв'язок (1) має вигляд лінійної залежності  $V(t) = V_0 + bt$ , яка відповідає регресійним кривим  $V_i(t_i)$  для часових рядів спостережень  $\{V_i\}_{i=1}^n$  на ранніх етапах зростання пухлини, коли обмеження ресурсів ще не впливає на динаміку зростання, або для спостережень на коротких відрізках часу.

Головні переваги лінійної моделі - це простота реалізації, прозорість інтерпретації параметрів та можливість швидко калібрувати її навіть на невеликих вибірках [11, 12]. Більшість пухлин на пізніх стадіях демонструють сповільнення зростання, що теж не описується лінійною моделлю [11, 36]. Однак, попри обмеження, у клінічній практиці лінійна регресія використовується для швидкої оцінки динаміки, попереднього скринінгу та коротких часових рядів. Складніші, змішані чи функціональні моделі дають змогу враховувати більш детальну структуру даних та індивідуальні особливості зростання [35, 37].

### 2.2 Поліноміальні моделі

На довгих проміжках часу пухлини демонструють нелінійне зростання об'єму і маси за рахунок більшої кількості клітин, що можна виявити з регресійних залежностей  $V(t)$ , які найкраще апроксимуються поліноміальними функціями  $V(t) = V_0 + \sum_{i=1}^n b_i t^i$ , де найчастіше  $i=2-4$ . Таким чином, зростання пухлини відповідає диференціальному рівнянню з початковою умовою

$$\frac{dV}{dt} = \sum_{i=1}^n i b_i t^{i-1}, \quad V(t_0) = V_0. \quad (2)$$

де коефіцієнти  $b_1, b_2$  мають зміст швидкості і прискорення процесу зростання.

Для біологічної інтерпретації рівняння (2) треба виявити загальний фізичний (і фізіологічний) зміст складників у правій частині (2), що важко зробити для  $n > 2$ . Крім того, зі зростанням  $n$  може погіршуватися стійкість моделі (2). У випадку  $n=2$  модель (2) добре вловлює зміну темпу розвитку пухлини та дозволяє фіксувати критичні точки (наприклад, момент переходу до плато чи вплив протипухлинної терапії). Поліноміальна регресія стала стандартним інструментом для аналізу даних у біостатистиці й часто слугує проміжною ланкою між простими та складнішими моделями (логістичною та Гомпертца). Зокрема, у змішаних моделях вона дозволяє моделювати як середню, так і індивідуальні траєкторії зростання [35].

### 2.3 Експоненціальна модель (модель Мальтуса)

Ця модель має прозору біологічну інтерпретацію, а саме швидкість зростання пропорційна масі (чисельності клітин) пухлини у даний момент часу, що відповідає біологічній тканині, усі клітини якої мають однакову ймовірність поділу, а зовнішні обмеження (ресурси, простір, імунна відповідь організму) ще не впливають на динаміку зростання. Історично ця модель була вперше запропонована англійським економістом Т. Мальтусом для зростання чисельності населення.

Математично експоненціальне зростання описується рівнянням:

$$\frac{dV}{dt} = bV + m, \quad V(t_0) = V_0, \quad (3)$$

де  $b$  – параметр Мальтуса,  $m$  – міграція клітин (переміщення від/до пухлини), розв’язок якого має вигляд  $V(t) = (V_0 + m/b)e^{b(t-t_0)} - m/b$ .

Таким чином, ця модель може використовуватися коли регресійний аналіз дає найкраще співвідношення до експоненціальної лінії тренду у даних часових рядів  $\{V_i\}_{i=1}^n$ . Серед основних переваг експоненціальної моделі — простота аналітичної форми, мінімальна кількість параметрів і легкість оцінки агресивності пухлини [12]. Вона ідеально підходить для ранніх фаз росту або для даних *in vitro*, коли вплив обмежуючих факторів мінімальний [11]. У популяційних та змішаних моделях параметр  $k$  дозволяє оцінити міжіндивідуальну варіабельність [35].

Втім, основне обмеження експоненціальної моделі — її непридатність для довгострокового прогнозування: у реальних біологічних системах ріст пухлин поступово уповільнюється через дефіцит кисню, поживних речовин та інші фактори [12,36]. В більшості випадків експоненціальна динаміка зберігається лише короткий час. У сучасних аналітичних системах експоненційну модель використовують як базову для порівняння складніших підходів або для попередньої оцінки параметрів.

#### 2.4. Експоненціальна модель, яка переходить у лінійну

У ряді робіт розглядається комбінування експоненціальної та лінійної динаміки для адаптації до особливостей реальних експериментальних даних, у тому числі зниження темпів зростання завдяки використанню хіміотерапії [34, 12, 38]. У цих моделях пухлина зростає експоненціально поки не досягне деякого критичного розміру  $V = V^*$ , після чого її зростання уповільнюється і апроксимується лінійною регресією, тобто

$$\frac{dV}{dt} = F(V) \equiv \begin{cases} \lambda V^* V, & V \leq V^* \\ \lambda, & V > V^* \end{cases}, \quad V(t_0) = V_0. \quad (4)$$

Відповідно до (4), перехід від експоненціального до лінійного зростання відбудеться у момент часу  $t^* = t_0 + \left(\lambda V^*\right)^{-1} \ln(\lambda V^* / V)$ .

Для обчислювальних цілей зручніше використовувати одне рівняння, яке описує перехід між фазами, особливо при врахуванні ефектів протипухлинних препаратів, що будуть введені в подальших дослідженнях. Тому в якості наближення використовується функція

$$F(w) = \lambda V^* V \left(1 + \left(V / V^*\right)^\kappa\right)^{-1/\kappa}, \quad \text{де параметр } \kappa \text{ визначає характер переходу між різними типами}$$

зростання. Ця функція узагальнює (3) і дозволяє моделювати плавний перехід між фазами зростання пухлини, зберігаючи можливість безперервного аналізу ефективності терапевтичних дій. Така модель добре відповідає зміні швидкості зростання різних типів пухлин під дією вдало підбраної терапії [34].

#### 2.5. Логістична модель Ферхюльста.

Логістична модель стала однією з найпопулярніших у біостатистиці, оскільки дозволяє враховувати обмеження ресурсів та ефекти насичення, які властиві реальним біологічним системам [35, 36]. З часом біологічним клітинам бракує поживних речовин і кисню, наростає гіпоксія та метаболічні обмеження, які уповільнюють зростання тканини.

$$\frac{dV}{dt} = bV \left(1 - \frac{V}{V_{\max}}\right), \quad V(t_0) = V_0, \quad (5)$$

де  $V_{\max}$  - максимальний можливий розмір (маса) пухлини,  $b$  – швидкість зростання.

Розв’язок (5) із заданою початковою умовою має вигляд S-подібної кривої

$$V(t) = \frac{V_{\max} V_0}{V_0 + (V_{\max} - V_0)e^{-bt}}, \quad (6)$$

яка описує як швидку початкову фазу зростання пухлини, так і період повільного зростання з виходом на плато.

Перевага логістичної моделі – у можливості роботи з індивідуальними та популяційними даними (змішані ефекти, врахування гетерогенності [35, 36]). Водночас модель не враховує складних біологічних механізмів (наприклад, гетерогенність та некроз пухлини, імунний вплив) і в деяких випадках поступається за точністю моделі Гомпертца [36]. Логістична модель забезпечує оптимальний баланс між простотою та точністю і залишається універсальним інструментом для математичного опису зростання пухлин у біостатистиці.

Важлива модифікація моделі (5) зв'язана з урахуванням ефекту Алле [10] і має вигляд [39]

$$\frac{dV}{dt} = bV \left(1 - \frac{V}{V_{\max}}\right) \left(1 - \frac{V}{V_{cr}}\right), \quad V(t_0) = V_0, \quad (7)$$

де  $0 < V_{cr} < V_{\max}$  - критичне значення, яке відповідає «переключенню» ракових клітин на «режим міграції» після досягнення критичного розміру пухлини.

Розв'язок (7) має вигляд

$$V(t) = \frac{V_{\max}(V_0 - V_{cr}) + V_{cr}(V_{\max} - V_0) \exp(bt(1 - V_{cr}/V_{\max}))}{V_0 - V_{cr} + (V_{\max} - V_0) \exp(bt(1 - V_{cr}/V_{\max}))}, \quad (8)$$

причому функція  $V(t)$  зростає при  $V \in ]V_{cr}, V_{\max}[$  і зменшується при  $V \in ]0, V_{cr}[$ .

Узагальнення моделі (5) на степеневу функцію у правій частині рівняння (узагальнена модель Річардса)

$$\frac{dV}{dt} = bV^\alpha \left(1 - \left(\frac{V}{V_{\max}}\right)^\beta\right)^\gamma, \quad V(t_0) = V_0, \quad (9)$$

в залежності від параметрів  $\alpha$  і  $\beta$  має вигляд від лінійного до S-подібних розв'язків різної крутизни, які відносяться до класу функцій Річардса.

Наявність у моделі (9) чотирьох параметрів дає можливість апроксимації експериментальних кривих з достатньо високою точністю [45]. При  $\alpha = \beta = \gamma = 1$  (9) переходить у класичну логістичну модель (5); при  $\alpha = \gamma = 1, \beta \neq 1$  отримуємо модель Річардса; при  $\alpha = 2/3, \beta = 1/3, \gamma = 1$  отримуємо логістичну модель у формі, яка була запропонована J.D. Murray [40] виходячи із загального закону збереження маси коли потік речовин до пухлини пропорційний площині її поверхні.

Оскільки клітинні цикли мають певну тривалість, синтез біохімічних компонентів, проліферація, міграція клітин та інші процеси впливають на подальші процеси зростання з деяким запізненням у часі  $\tau$ , замість (5) розглядається рівняння із запізненням у часі [41, 42]

$$\frac{dV(t)}{dt} = bV(t-\tau) \left(1 - \frac{V(t-\tau)}{V_{\max}}\right). \quad (10)$$

## 2.6. Функція Бергаланфі

Модель відповідає припущенню, що на клітинному рівні біологічне зростання є результатом балансу між анаболізмом і катаболізмом, тобто між синтезом і розпадом біологічних молекул, швидкості яких є степеневими функціями маси клітин. Таким чином, балансове рівняння зростання має вигляд:

$$\frac{dV}{dt} = bV^n - aV^m, \quad V(0) = V_0, \quad (11)$$

де  $a, b, n, m$  – параметри моделі, які залежать від використання ресурсів, метаболізму та морфологічної структури пухлини.

Розв'язок (11) у загальному вигляді є  $w(t) = w_0 + \int_0^t \frac{dt}{bw^n - aw^m}$ . Біологічно значущим є випадок

$n=1, m \leq 1$ , для якого розв'язок можна обчислити шляхом інтегрування у вигляді

$V(t) = \left( \frac{a}{b} + \left( V_0^{1-m} - \frac{a}{b} \right) e^{b(1-m)t} \right)^{\frac{1}{1-m}}$ . Наявність чотирьох параметрів у моделі (5) дає можливість

гнучко апроксимувати експериментальні дані функцією Берталанфі, яка дає найкращі результати у випадку пухлин різних типів [43, 44, 45]. Також, по аналогії з (10), розглядаються рівняння Берталанфі (11) із запізненнями у часі.

## 2.7. Модель Гомпертца

Ця універсальна модель використовується для моделювання зростання чисельності популяцій мікроорганізмів, розмірів тіла тварин та окремих частин рослин, у тому числі пухлин, для аналізу динаміки смертності у демографічних дослідженнях та у моделюванні економічних процесів. Диференціальне рівняння моделі із початковою умовою

$$\frac{dV}{dt} = bV e^{-ct}, \quad V(0) = V_0 \quad (12)$$

має розв'язок  $V(t) = V_0 \exp\left(\frac{b}{c}(1 - \exp(-ct))\right)$ ,  $b > 0, 0 < c < 1$ , де  $c$  – швидкість затухання проліферації.

Функція Гомпертца відтворює поступове уповільнення росту до появи насичення, а точка перегину функції розташована несиметрично, чим відрізняється від логістичної моделі [36]. Саме це дозволяє якісно моделювати дані для середніх і великих пухлин та уникати переоцінки об'єму у довгостроковому прогнозі [12]. З біологічної точки зору, гальмування зростання пов'язують з виснаженням ресурсів мікрооточення, наростанням гіпоксії, метаболічними зрушеннями і реакцією імунної системи організму [7]. Порівняльні дослідження показують, що саме крива Гомпертца для багатьох пухлин (гліоми, карциноми, раку грудей, моделі *in vivo*) часто дає найкраще узгодження з реальними траєкторіями зростання, особливо для тривалих періодів спостереження [11, 36, 46, 25].

Усі обговорені вище моделі (1) - (5), (7), (9), (11), (12) задовольняють отриманим експериментально ростовим кривим  $V(t)$  за рахунок підбору параметрів [10] і не існує однієї «найкращої» моделі [47]. Різні типи пухлин відповідають різним моделям, а також існують індивідуальні параметри, які пов'язані з особливостями організму пацієнта. Така невизначеність зворотних залежностей між динамікою зростання та конкретною математичною моделлю стимулює застосування сучасних алгоритмів, які використовують машинне навчання (Machine Learning, ML) і, особливо, глибоке машинне навчання (Deep Machine Learning, DML).

## 2.8. Статистичні методи аналізу даних

Статистичні методи аналізу даних спостережень та експериментальних вимірювань динаміки зростання пухлин у вигляді часових рядів широко використовуються для перевірки наявності закономірностей, кореляцій і т. д. для вибору однієї з математичних моделей (1)-(12), або їх комбінацій чи модифікацій для прогнозування динаміки онкологічної хвороби та її лікування. Для цього використовуються методи дескриптивної (описової) статистики, регресійний, кореляційний і факторний аналіз [47, 49]. Попередньо часові ряди мають бути згладжені (фільтрація випадкового шуму), перевірені на наявність пропущених даних та викидів, які пов'язані із недосконалістю процесу вимірювання. Як відомо, «сирі» медичні дані (raw data) рідко відповідають вимогам математичного моделювання. Вони можуть містити пропуски, аномалії, мати різні одиниці вимірювання, широкий діапазон варіацій та інші особливості, що суттєво ускладнює аналіз. Якісна попередня обробка сирих даних є необхідною умовою для подальшого застосування математичних моделей [35, 38].

### 2.8.1. Нормалізація даних

Нормалізація даних з приведенням до безрозмірних величин шляхом ділення на максимальне значення вимірюваної величини у часовому ряді, а також перевірка нормального (Гаусового) розподілу є важливою для застосування класичних статистичних методів. В результаті отримуємо часові ряди зі значеннями вимірюваних параметрів у діапазоні  $V_i \in [0, 1]$ .

### 2.8.2. Викиди.

Медичні дані часто містять нетипові або помилкові значення, які можуть суттєво вплинути на результати моделювання. Для їх виявлення використовують як класичні статистичні методи -

аналіз boxplot, Z-score, IQR (інтерквартильний розмах), так і алгоритми ML. Викиди або виправляють, або виключають із вибірки залежно від контексту [11, 38].

### 2.8.3. Трансформації даних.

Для наближення розподілу до нормального часто застосовують логарифмічне, кореневе чи Вох-Сох перетворення. Такі операції дозволяють підвищити ефективність регресійних моделей, особливо у випадках із сильними перекосами розподілу даних відносно Гаусового [35].

### 2.8.4. Імпутація пропусків.

У медичних наборах даних пропуски трапляються часто (наприклад, через нерегулярні обстеження, людський фактор або технічні збої). Для їх заповнення використовують прості підходи (середнє, медіана, останнє спостереження) або сучасні алгоритми, такі як метод k-найближчих сусідів (k-nearest neighbors, KNN), випадковий ліс (Random Forest, RF) мультиімпутація та інші [35, 38]. Це дозволяє уникнути втрати цінної інформації та підвищити статистичну достовірність аналізу.

### 2.8.5. Зниження розмірності.

При роботі з великим числом ознак (наприклад, різних характеристик з медичних зображень пухлин) актуальними є методи відбору (feature selection) та зниження розмірності датасету за допомогою аналізу головних компонент (Principal Component Analysis, PCA), алгоритму розподіленого стохастичного вбудовування сусідів (t-distributed stochastic neighbor embedding, t-SNE), та ряду інших. Це дозволяє зосередитися на найбільш інформативних параметрах та уникнути перенавчання.

Комплексне застосування зазначених статистичних процедур забезпечує максимальну якість і достовірність аналізу, дозволяє працювати з великими, гетерогенними, частково неповними наборами даних — тобто в умовах, які характерні для сучасної біоінформатики та клінічної онкології [12, 38].

### 2.8.6. Регресійний і кореляційний аналіз.

Різні типи ліній регресії (лінійна, поліноміальна, степенева, експоненціальна, логарифмічна, S-подібна) будувалися методом найменших квадратів [49]. Найкраща значуща регресія обрала відповідно до максимального значення коефіцієнта детермінації

$$R^2 = 1 - \frac{\sum_{i=1}^n (V_i - V_i^{\text{mod}})^2}{\sum_{i=1}^n (V_i - \bar{V}_i)^2}, \quad (13)$$

де  $\{V_i\}_{i=1}^n$  - виміряні часові ряди,  $\bar{V}_i$  - середнє значення ряду,  $V_i^{\text{mod}}$  - значення, прогнозовані за обраною математичною моделлю.

У залежності від отриманої регресійної кривої, обирається відповідна математична модель (лінійного, поліноміального, експоненціального та ін. зростання).

Кореляції між двома кривими зростання пухлин і обчислювалась на основі коефіцієнта кореляції Пірсона

$$K_p = \frac{\sum_{i=1}^n (X_i - M(X))(Y_i - M(Y))}{\sqrt{\sum_{i=1}^n (X_i - M(X))^2 (Y_i - M(Y))^2}}, \quad (14)$$

де  $\{X_i\}_{i=1}^n$  і  $\{Y_i\}_{i=1}^n$  - часові ряди,  $M(X_i)$ ,  $M(Y_i)$  - їх середні значення (математичні очікування).

### 2.8.7. Критерії якості обраної математичної моделі.

Для порівняння різних масивів даних потрібні обчислення середніх значень (математичне очікування, медіана, мода) та розкиду даних (дисперсія або стандартне відхилення). Для перевірки точності математичної моделі найчастіше використовують критерій  $R^2$  (13). Статистично значущими є значення  $R^2 > 0.8$ , але у біомедичних даних часто використовують невеликі датасети даних, які отримані на групах із 20-40 хворих, і тому статистичні методи, строго кажучи, не можуть бути використані, а значення  $R^2 > 0.5$  вважаються за істотними.

Для кількісної оцінки точності тієї чи іншої математичної моделі використовуються також середня абсолютна помилка (Mean Absolute Error, MAE)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |V_i - V_i^{\text{mod}}|, \quad (15)$$

або середньоквадратична помилка (Root Mean Squared Error, RMSE)

$$\text{RMSE} = \left( \frac{1}{n} \sum_{i=1}^n (V_i - V_i^{\text{mod}})^2 \right)^{1/2}, \quad (16)$$

які чутливі до великих відхилень і добре підходять для датасетів із рідкими, але значними аномаліями [12].

Крім основних статистичних показників (13), (15), (16), для порівняння математичних моделей із різною складністю використовуються інформаційні критерії, такі як Akaike Information Criterion (AIC) та Bayesian Information Criterion (BIC). Вони дозволяють визначати, чи виправдане збільшення числа параметрів поліпшенням точності, і допомагають уникати перенавчання алгоритмів ML. Додатково для оцінки нормальності розподілу залишків може застосовуватись тест Колмогорова-Смірнова чи аналіз автокореляції залишків тестом Дарбіна-Вотсона [12, 36].

Завдяки цим інструментам стає можливим обирати не лише модель із найкращою підгонкою, а й ту, що має оптимальний баланс між точністю, складністю і стійкістю до шуму у даних. Таким чином, стандартні статистичні методи потребують ретельної підготовки сирих даних до аналізу, що займає значний час. При цьому сучасні методи ML, DML можуть ефективно обробляти сирі дані, що значно зберігає час.

## 2.9. Підхід Random Forest (RF)

RF - це ансамблевий метод машинного навчання, який дозволяє ефективно моделювати складні нелінійні залежності у біомедичних задачах, зокрема для аналізу динаміки зростання пухлин [38, 50]. Головна ідея методу - це побудова ансамблю дерев рішень, де кожне дерево тренується на випадковий підвбірці даних і підмножині ознак, а прогноз визначається як середнє усіх дерев. Така архітектура забезпечує стійкість до викидів, автоматичну обробку пропусків та зниження ризику перенавчання моделі. Порівняно з класичними математичними моделями (1)-(12), RF не потребує визначення конкретної функціональної форми й може апроксимувати навіть складні S-подібні чи багатозначні криві зростання. Особливо актуально це для тих задач, у яких дані вимірювань містять шум, а структура залежностей між даними заздалегідь невідома [51].

Серед ключових переваг RF — гнучкість, висока точність прогнозування, можливість роботи з великою кількістю предикторів і ефективна інтеграція різних типів ознак (радіомічних, генетичних, клінічних). У клінічній онкології метод використовується для індивідуального прогнозування траєкторій зростання пухлини, аналізу факторів ризику та виживаності, а також для створення цифрових двійників пацієнтів (digital twin) [38, 50]. Недоліки RF — відсутність явної біологічної інтерпретації параметрів, складність аналізу впливу окремих змінних і потреба у достатньо великій вибірці для досягнення стабільної точності моделі. Особливий розвиток отримали Випадкові ліси виживання (Random Survival Forests) — модифікація RF для аналізу часу до певної події, що дуже корисно у медичних дослідженнях [50].

Таким чином, комбінація класичних статистичних методів та сучасних алгоритмів підготовки і обробки даних вимірювань дозволяють знайти найбільш адекватну математичну модель, провести її валідацію на датасетах медичної інформації та використовувати для прогнозу динаміки зростання пухлини та вибору персоналізованого лікування. Математичні моделі (1)-(12) у вигляді одного чи системи ЗДР можуть бути розширені на неоднорідні пухлини з їх мікрооточенням шляхом додавання відповідних дифузійних членів у рівняння, що дає системи ДРЧП для більш детального опису просторово-часових процесів, які пов'язані із розвиненням пухлини та змінами оточуючих її тканин.

## 3. Матеріали і методи.

### 3.1. Джерело інформації.

У цій роботі виконаний порівняльний аналіз статистичних та сучасних машинних методів на прикладі гліобластоми (GBM), дані зростання якої отримано з набору даних LUMIERE (Longitudinal Glioblastoma MRI with expert RANO evaluation), яка міститься у відкритому джерелі [52].

База даних охоплює 91 пацієнта з діагнозом гліобластома, для яких проведено 638 МРТ-обстежень (загалом 2487 окремих зображень). Кожне обстеження включає мультипараметричні МРТ-зображення, отримані в основних режимах: T1 (до введення контрасту), T1c/T1Gd (після введення контрасту), T2w та FLAIR (Fluid-Attenuated Inversion Recovery). Така комбінація забезпечує детальну візуалізацію структури пухлини та прилеглих тканин. До набору додаються результати автоматичної сегментації пухлинних зон, які надаються разом із МРТ-зображеннями.

Тут виокремлюють три ключові компоненти пухлини:

1. Некротична/неактивна зона NEC (Necrotic/NonEnhancing core): центральна частина пухлини, яка часто складається з відмерлих клітин і не накопичує контрастну речовину.
2. Ділянка з контрастним підсиленням CEL (Enhancing Core/Contrast Enhancing Lesion): активна частина пухлини, що інтенсивно накопичує контраст і свідчить про агресивне зростання.
3. Зона набряку PEC (Edema/Peritumoral Edema Compartment): область навколо пухлини, що характеризується накопиченням рідини внаслідок запальних процесів або порушення відтоку.

У супровідних JSON-файлах (JavaScript Object Notation) подано об'єми кожного з компонентів пухлини, що дозволяє проводити аналіз її змін з часом. Додатково, у наборі є дані про пацієнтів: стать, вік, особливості гена MGMT, наявність мутації в гені IDH1, тривалість життя після лікування та результати терапії за шкалою RANO (Response Assessment in Neuro-Oncology). Для цілей даного дослідження було обрано підмножину з 599 випадків, у яких представлені всі три компоненти пухлини, що є репрезентативною вибіркою для статистично достовірного аналізу даних. Багатовимірний характер бази даних LUMIERE, яка поєднує зображення, кількісні параметри та клінічні характеристики, створює базу для комплексного аналізу динаміки гліобластоми з урахуванням індивідуальних особливостей пацієнтів.

### 3.2. Попередня обробка та структурування даних для аналізу.

Початкові дані були збережені у форматі JSON та містили числові показники об'ємів сегментів пухлини. З метою їх подальшого аналізу ці дані були імпортовані й упорядковані у формат табличних структур із використанням мови програмування Python та бібліотеки pandas. Основну увагу було приділено змінним, що характеризують об'єми трьох ключових компонентів пухлини: NEC, CEL, і PEC для всіх доступних моментів спостереження у часі, що вимірювався у тижнях. У подальшому ці дані були об'єднані з інформацією із додаткових таблиць, що включали:

- Демографічну інформацію: вік в якому було проведено хірургічне втручання (Age at surgery (years)) та стать (Sex) пацієнтів;
- Клінічні змінні: статус метилування промотора гена MGMT (MGMT qualitative), мутаційний статус IDH (IDH (WT: wild type), IDH method);
- Категорії відповіді за критеріями RANO (Rating);
- Загальний час виживання (LessThan3Months як додаткова змінна).

Інтеграція здійснювалася шляхом злиття таблиць за спільними колонками за унікальним ідентифікатором пацієнта (Patient) та відповідною часовою точкою (Wheel). На наступному етапі було проведено очистку даних. З метою забезпечення коректності математичних операцій усі числові поля було приведено до відповідних типів (float, int), а категоріальні – до уніфікованих ярликів. Назви колонок було стандартизовано (видалено спеціальні символи, замінено пропуски на підкреслення тощо) для підвищення зручності роботи з даними.

У результаті попередньої обробки був сформований структурований дата-фрейм, який включає 599 записів (по одному на кожне спостереження), а також, залежно від наявності допоміжної інформації, 10–12 змінних, а саме:

1. Patient\_N: Унікальний ідентифікатор пацієнта (N=1, 2,...,91).
2. Week: Інтервал часу у тижнях відносно першого дослідження.
3. NEC, CEL, PEC: Об'єми відповідних сегментів пухлини (мм<sup>3</sup>, тип float).
4. Age (роки): Вік пацієнта на момент операції (тип int).
5. Sex: Стать пацієнта ('male' або 'female').
6. IDH (WT: wild type): Мутаційний статус IDH ('Mutated', 'WT', 'Unknown').

7. IDH method: Метод визначення статусу IDH.
8. MGMT qualitative: Метилування промотора MGMT (якісна і кількісна оцінка).
9. Rating: Категорія відповіді за критеріями RANO (‘Complete Response’, ‘Partial’, ‘Stable’, ‘Progressive’).
10. LessThan3Months: Двійкова змінна, яка вказує на виживання менше 3 місяців після операції (0 або 1).
11. Date: Дата МРТ-дослідження.
12. NonMeasurableLesions: Додаткова характеристика уражень, що не підлягають вимірюванню.

У досліджуваній групі середній вік становив  $\sim 60,7 \pm 14,6$  років. Розподіл за статтю був майже рівномірним: 44 жінки та 47 чоловіків. Статус метилування MGMT був відомий для 80 пацієнтів: 37 мали метильований промотор, 43 – неметильований, ще 11 – невідомий. Статус IDH1 був відомий для 58 пацієнтів, серед яких лише один мав виявлену мутацію (решта – WT). Через надто малу чисельність групи з IDH1-мутацією порівняльний аналіз за цим маркером не проводився.

### 3.3. Результати статистичного аналізу датасетів.

Для кожного компонента пухлини обчислено базові статистичні характеристики: середнє значення, медіана, стандартне відхилення та діапазон значень (Таблиця 1). Отримані результати свідчать про суттєву міжіндивідуальну варіативність об’ємів пухлин. Для виявлення потенційних залежностей між об’ємами компонент проведено кореляційний аналіз із використанням коефіцієнту Пірсона (14). Найвищу кореляцію виявлено між CEL і PEC ( $R^2 = 0.54$ ), що свідчить про помірний взаємозв’язок між зростанням активної частини пухлини та перифокальним набряком. Кореляція між NEC та іншими компонентами є низькою ( $R^2 = 0,14-0,33$ ). Відповідна матриця кореляцій наведена на Рис.1.

Таблиця 1. Середні значення для різних компонентів гліобластоми (GBL)  
Table 1. Average values for different components of glioblastoma (GBL)

| Компонент | Діапазон (мм <sup>3</sup> ) | Середнє (мм <sup>3</sup> ) | Медіана (мм <sup>3</sup> ) | Дисперсія (мм <sup>3</sup> ) |
|-----------|-----------------------------|----------------------------|----------------------------|------------------------------|
| NEC       | 0-150                       | 25.0                       | 15.0                       | 30.0                         |
| CEL       | 0-300                       | 40.8                       | 30.0                       | 35.0                         |
| PEC       | 5-450                       | 62.7                       | 45.0                       | 50.0                         |

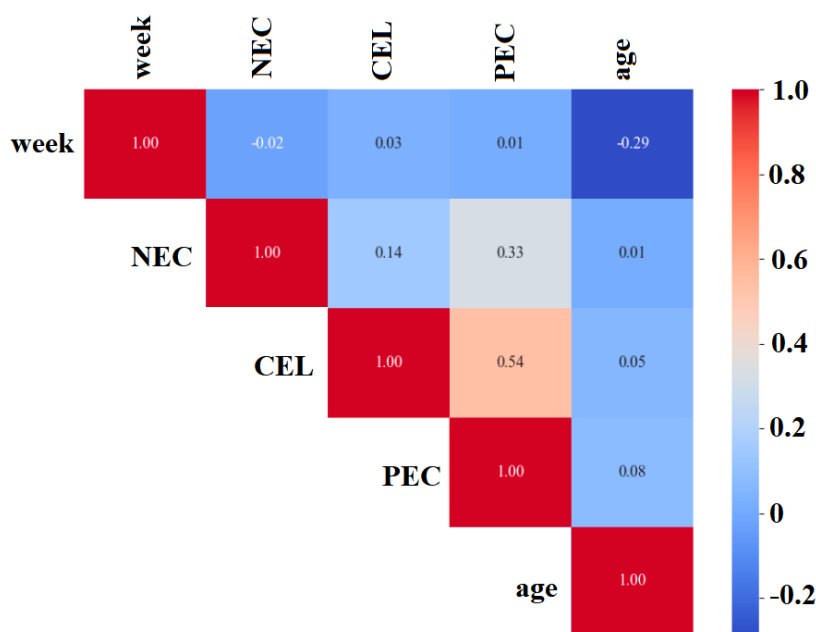


Рис.1. Кореляційна матриця числових ознак.  
Fig. 1. Correlation matrix of numerical features.

Додатково проведено порівняння обсягів компонент між підгрупами пацієнтів, зокрема за статусом MGMT та статтю. Найбільш виражені відмінності спостерігались між групами MGMT: пацієнти з метильованим промотором мали статистично значущо менші значення Enhancing\_Core порівняно з неметильованими ( $p < 0,05$ ), що узгоджується з попередніми клінічними даними про кращий прогноз і менш агресивний ріст при MGMT-метилації.

Щодо статі пацієнта, розподіл обсягів CEL між чоловіками та жінками був подібним: медіанні значення та інтерквартильний розмах практично не відрізнялись, а кількість викидів була приблизно однаковою в обох групах. Це узгоджується з літературними даними про відсутність істотного гендерного впливу на макроструктурні характеристики GBM [47].

Аналіз впливу віку на об'єм CEL продемонстрував, що різниця між молодшою (<60 років) та старшою ( $\geq 60$  років) віковими групами статистично незначуща (U-критерій Манна-Уїтні,  $p = 0.787$ ). Незважаючи на певну тенденцію до вищих медіанних значень у старшій групі, результат не дозволяє стверджувати про суттєвий вплив віку на розвиток перифокального набряку у пацієнтів цієї групи.

Окрім порівняльного аналізу, було досліджено динаміку змін об'ємів пухлин у часі для окремих пацієнтів. У якості прикладу, на Рис.2 наведені часові ряди об'ємів трьох компонент пухлини для пацієнта Patient-025. Помітно поступове нелінійне зростання CEL та PEC. Середні об'єми компонент пухлин та їх розкид за даними 599 випадків наведені на Рис.3. Як видно з рисунка, компонент PEC характеризується найбільшими об'ємами та високою варіабельністю. Натомість об'єми NNE і CEL є меншими, хоча теж мають значну кількість викидів.

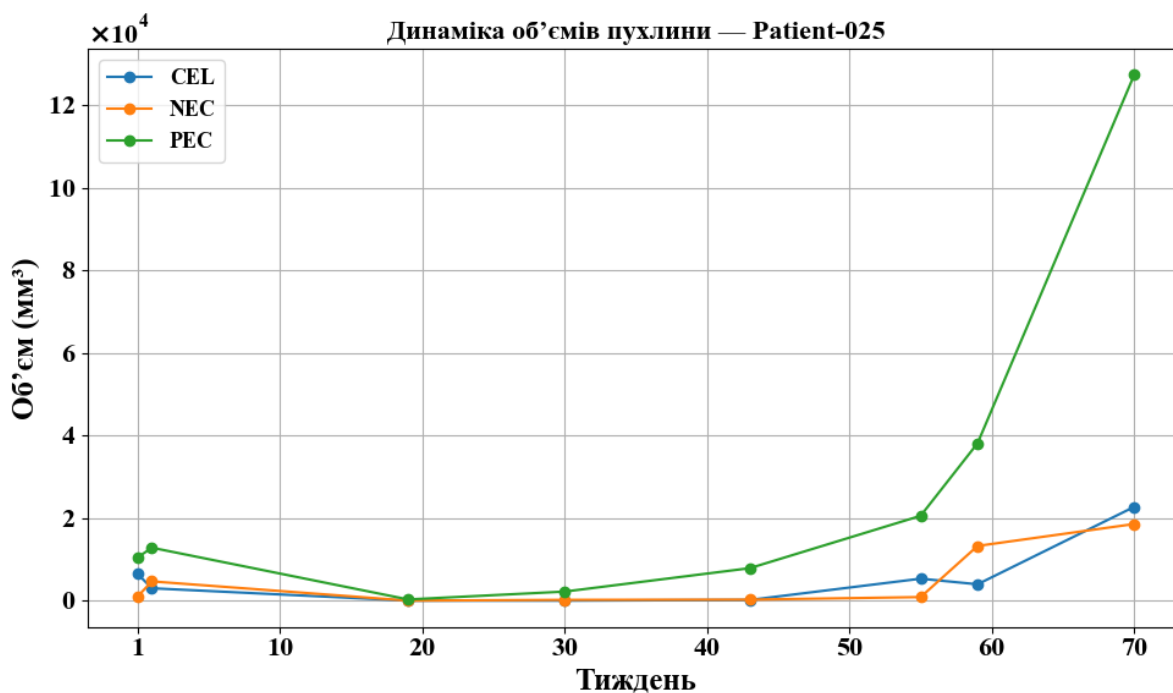


Рис.2. Динаміка у часі об'ємів NEC, CEL і PEC пухлини для пацієнта Patient-025.

Fig. 2. Dynamics of NEC, CEL, and PEC tumor volumes over time for Patient-025.

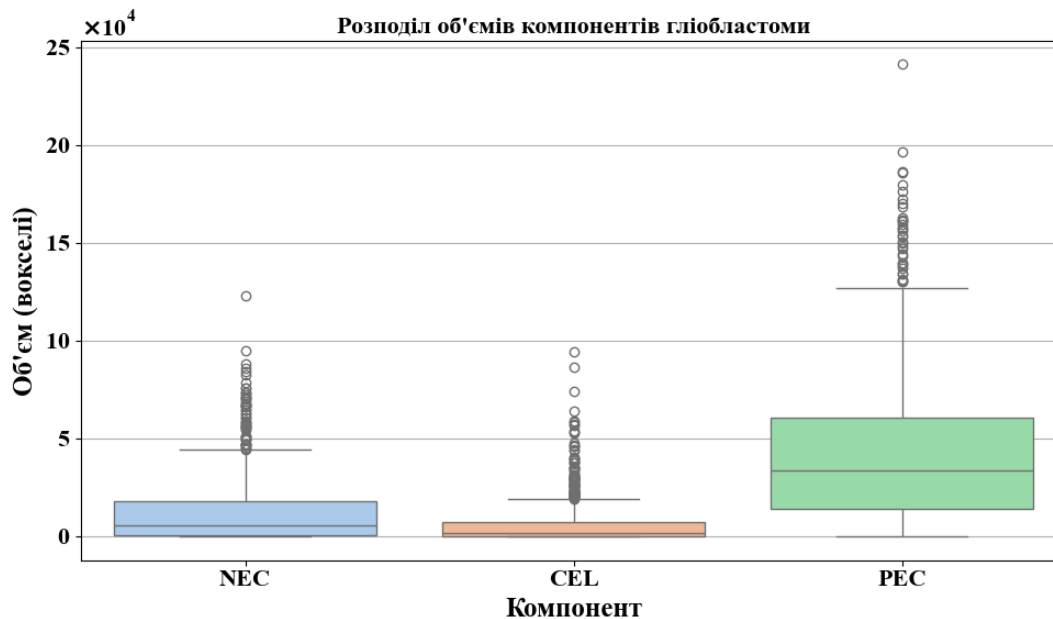


Рис.3. Розподіл об'ємів компонентів гліобластоми: NEC, CEL, PEC.

Fig. 3. Distribution of glioblastoma component volumes: NEC, CEL, PEC.

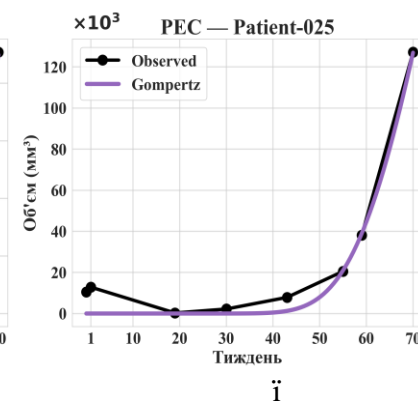
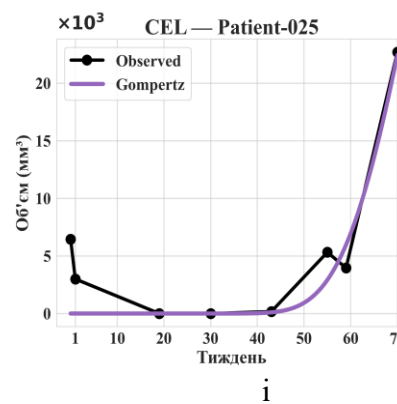
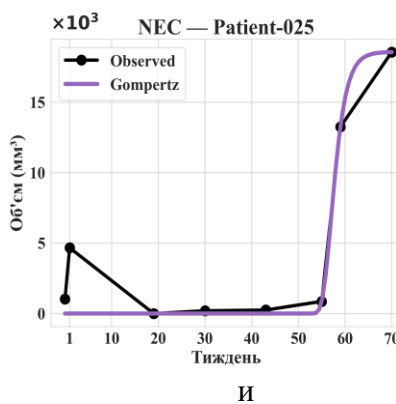
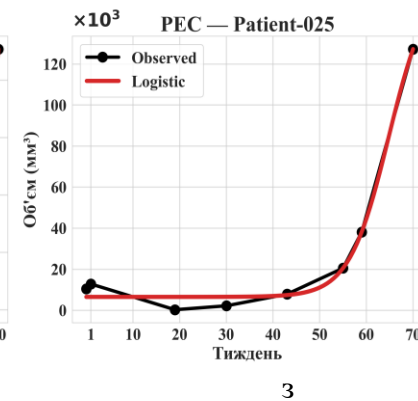
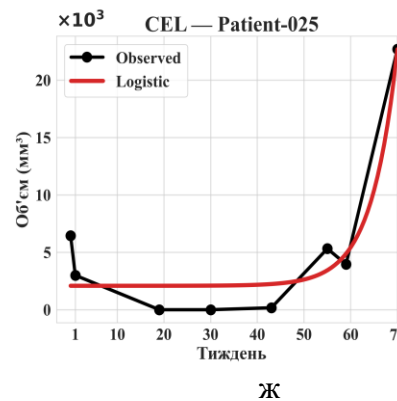
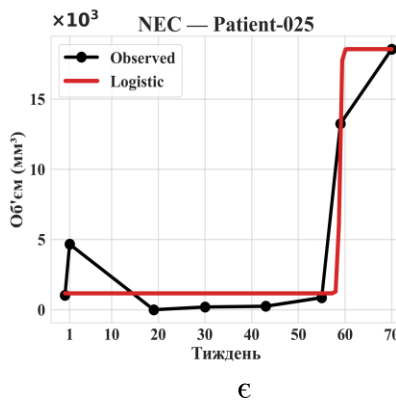
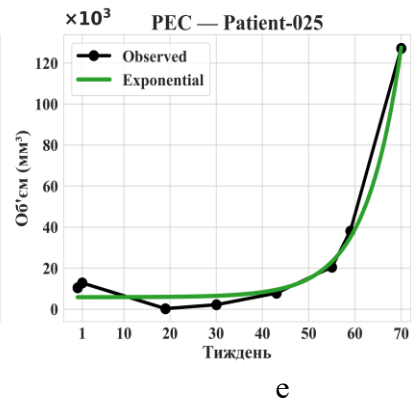
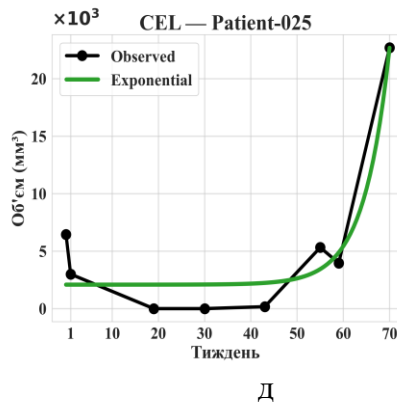
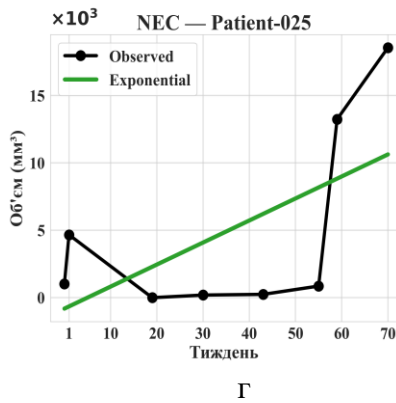
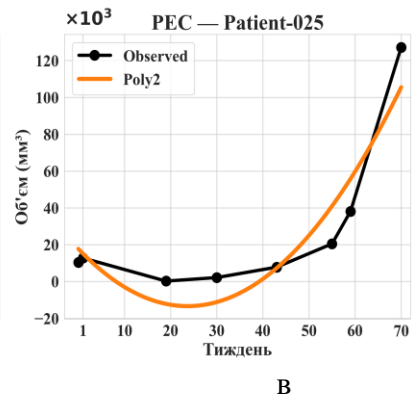
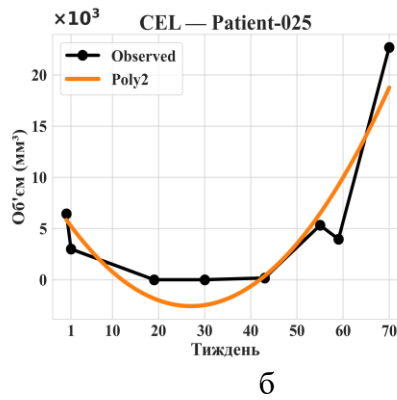
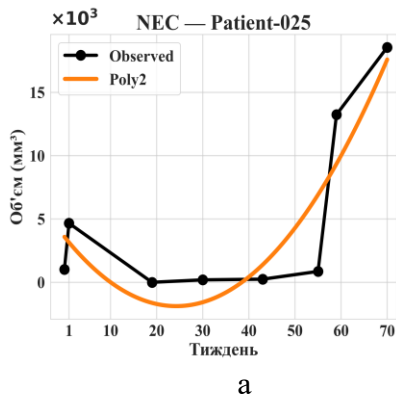
Загалом, результати описової статистики, порівняльного та кореляційного аналізу дозволяють сформувати цілісне уявлення про варіативність об'ємів компонентів GBM у групі пацієнтів. Виявлені зв'язки (зокрема між об'ємами активного ядра та набряку) і тенденції (менші об'єми при MGMT-метилації) можуть бути використані для параметризації математичних моделей росту пухлини. Додатково, наявність повних часових рядів створює передумови для побудови індивідуалізованих моделей динаміки зростання пухлини.

#### 4. Математичне моделювання зростання пухлини.

##### 4.1. Регресійний аналіз.

Регресійний аналіз індивідуальних кривих зростання пухлин пацієнтів показав відсутність лінійної регресії ( $R^2 \sim 0.1-0.3$ ), більш значуще наближення поліномами 2-го порядку ( $R^2 \sim 0.37-0.72$ ) та експонентною ( $R^2 \sim 0.224-0.623$ ). Експоненціальна модель виявилася ефективною для окремих траєкторій із вираженим початковим зростанням, проте для тривалих або складних кривих її прогностичні властивості недостатні. Наближення S-подібними кривими показало значущі результати для логістичної кривої ( $R^2 \sim 0.371-0.690$ ) і невисокі результати – для функції Гомпертца ( $R^2 \sim 0.012-0.029$ ).

Метод RF з 200 деревами і обмеженням глибини ( $\max\_depth=3$ ) був застосований для кривих зростання кожної з 3-х частин пухлини кожного з 599 хворих. Така конфігурація була обрана для запобігання перенавчанню, враховуючи обмежену кількість точок для кожного випадку. На кожному кроці фіксувалися значення метрик якості  $R^2$ , RMSE, MAE, MAPE а також параметри моделі. У якості приклада результати регресійного наближення кривих зростання об'ємів NEC, CEL, PEC у даних Patient-025 наведені на Рис.4. Середні значення якості наближень (значення  $R^2$  та RMSE) наведені на Рис.5. Таким чином, динаміка зростання CEL характеризується помірно високими значеннями  $R^2 \sim 0.37$  і мінімальним значенням RMSE. Криві зростання PEC мають найбільші значення  $R^2$ , але значно більший розкид. Динаміка зростання NEC слабо піддається використанню регресійним моделям за рахунок раптових змін у швидкості зростання з прискореннями та уповільненнями.



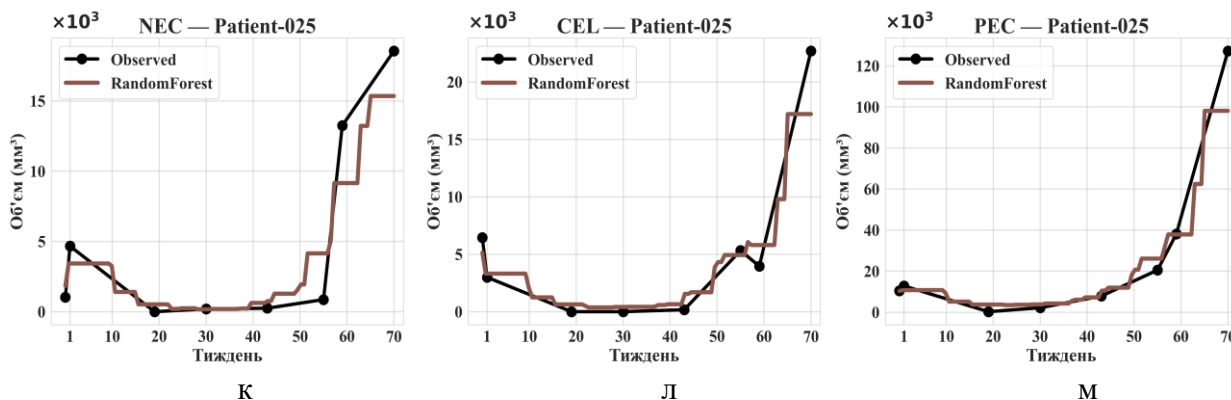


Рис.4. Результати апроксимації результатів вимірювання кривих зростання об'ємів NEC (а,г,е,и,к), CEL (б,д,ж,і,л), PEC (в,е,з,ї,м) параболічною (а-в), експоненціальною (г-е), логістичною (є-з), гомперцевою (и-ї) та RF (к-м) функціями для пацієнта Patient-025.

Fig. 4. Results of approximation of measurement results for NEC (a, d, e, f, g) and CEL (b, c, d, e, f) growth curves, PEC (c, e, g, h, m) growth curves using parabolic (a-c), exponential (d-e), logistic (f-g), Gompertz (h-i) and RF (k-m) functions for patient Patient-025.

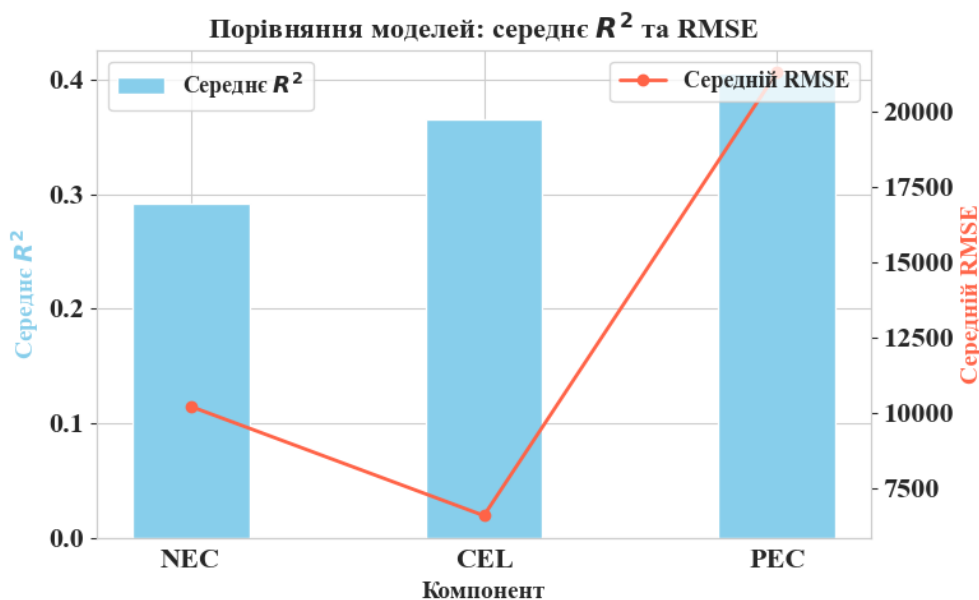


Рис.5. Порівняння середніх значень R<sup>2</sup> та RMSE для кривих зростання NEC, CEL, PEC для Patient-025.  
 Fig. 5. Comparison of average R<sup>2</sup> and RMSE values for NEC, CEL, and PEC growth curves for Patient-025.

#### 4.2. Прогнозування динаміки зростання пухлин.

Після визначення найкращого (у сенсі параметрів R<sup>2</sup>, MAE, RMSE та ін.) наближення (регресії) до кривої зростання відповідної області пухлини (NEC, CEL, PEC) у даного пацієнта, можна прогнозувати динаміку зростання шляхом інтерполяції. У роботі було здійснено прогноз об'єму пухлини на 8 тижнів наперед, що відповідає типовому проміжку між контрольними обстеженнями у клінічній практиці. На основі моделі найкращого наближення можна провести екстраполяцію кривої від останніх точок (59 та 70 тижнів) до наступної (78 тижнів) відповідно до моделі. На Рис.6 наведений приклад прогнозування для однієї з кривих. Крім того, можна провести ідентифікацію параметрів відповідної математичної моделі (2)-(5), (7), (9), (11) або (12) та використовувати її розв'язок із відомими коефіцієнтами для прогнозування як подальшого зростання пухлини, так і можливих наслідків її лікування різними препаратами. Для цього потрібні лабораторні дані про чутливість даної пухлини до ліків різних типів. Відповідний доданок треба ввести до правої частини ЗДР у якості негативного джерела для V(t) у вигляді лінійної або періодичної функції від дози препарату з урахуванням поглинання та розкладу препарату в організмі з часом, відповідно до кінетики Міхаеліс-Менетен або іншої.

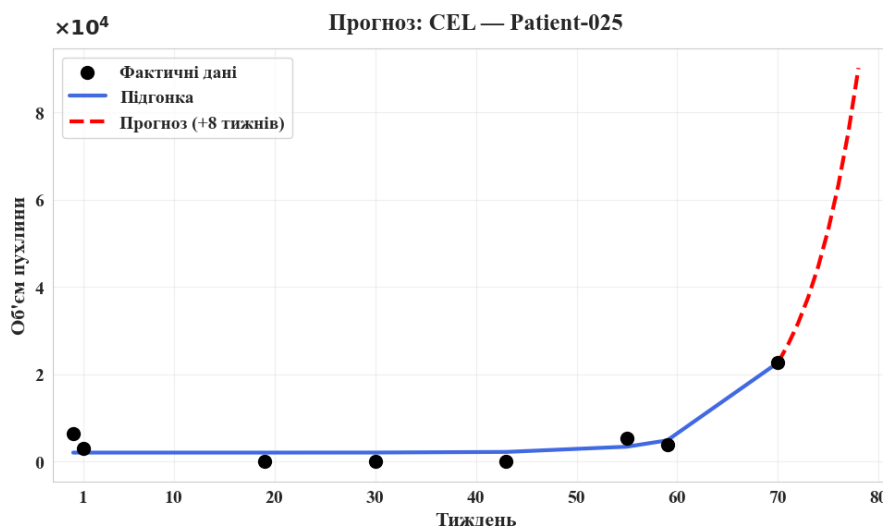


Рис.6. Прогноз динаміки об'єму компоненти CEL у пацієнта Patient-025.

Fig. 6. Forecast of the dynamics of the CEL component volume in patient Patient-025.

## 6. Висновки.

Результати роботи підтвердили практичну цінність математичного моделювання для оцінки динаміки зростання пухлин, можливостей оптимізації терапії та прогнозування результатів лікування. Поєднання аналітичних моделей із алгоритмами ML, разом із якісною підготовкою даних (нормалізація, імпуція пропусків, згладження, перетворення ознак) відкриває перспективи для подальшого розвитку персоналізованих підходів у сучасній медичній практиці та біостатистиці.

На основі аналізу одного клінічного випадку проведено порівняння класичних математичних моделей зростання пухлини (лінійна, поліноміальна, експоненційна, логістична, Берталанфі та Гомпертца), а також сучасних інструментів ML, зокрема RF. Кожна з моделей продемонструвала свої сильні та слабкі сторони: лінійна і експоненціальна моделі добре описують початкові фази, логістична та Гомпертцева — динаміку із фазою насичення, а RF забезпечує адаптацію до складних, гетерогенних і частково неповних даних, однак має слабку інтерпретованість.

Виявлено, що для тривалих та насичених спостережень найкраще працюють нелінійні моделі, тоді як короткі часові ряди доцільно аналізувати за допомогою лінійної чи поліноміальної регресії, для яких є прозора біологічна інтерпретація. Використання сучасних алгоритмів ML значно розширює можливості кількісного аналізу пухлинного росту, особливо у випадках складної структури даних чи відсутності чіткої аналітичної форми моделі.

Саме комбінований підхід - об'єднання класичних та сучасних статистичних інструментів - дає змогу отримати гнучкі та інтерпретовані рішення для персоналізованої медицини. Результати можуть стати основою для впровадження цифрових двійників (digital twin) пацієнта, підвищення точності діагностики і розробки нових індивідуалізованих стратегій у клінічній онкології.

## REFERENCES

1. World Health Organization, Cancer, 3 February 2025, [Online] Available: <https://www.who.int/news-room/fact-sheets/detail/cancer> Accessed on: February 01, 2026.
2. R.L. Siegel, T.B. Kratzer, A.N. Giaquinto, et al., "Cancer statistics, 2025", CA: A Cancer Journal for Clinicians, 75(1), 10–45, 2025. doi:10.3322/caac.21871
3. A.O. Lawal, T.J. Ogunniyi, O.I. Oludele, et al., "Innovative laboratory techniques shaping cancer diagnosis and treatment in developing countries", Discover Oncology, 16, 137, 2025. doi:10.1007/s12672-025-01877-w
4. G. Molla, M. Bitew, "The Future of Cancer Diagnosis and Treatment: Unlocking the Power of Biomarkers and Personalized Molecular-Targeted Therapies", Journal of Molecular Pathology, 6, 20, 2025. doi:10.3390/jmp6030020

5. H.K. Elaibi, F.F. Mutlag, E. Halvaci, et al., “Comparison of Traditional and Modern Diagnostic Methods in Breast Cancer”, *Measurement*, 242, 116258, 2025. doi:[10.1016/j.measurement.2024.116258](https://doi.org/10.1016/j.measurement.2024.116258)
6. R.P. Araujo, D.L.S. McElwain, “A history of the study of solid tumour growth: the contribution of mathematical modelling”, *Bulletin of Mathematical Biology*, 66(5), 1039–1091, 2004. doi: 10.1016/j.bulm.2004.03.001
7. D. Hanahan, R.A. Weinberg, “The Hallmarks of Cancer”, *Cell*, 100(1), 57–70, 2000. doi:10.1016/S0092-8674(00)81683-9
8. J. Folkman, “Tumor Angiogenesis: Therapeutic Implications”, *The New England Journal of Medicine*, 285, 1182–1186, 1971. doi:10.1056/NEJM197111182852108
9. P. Carmeliet, R.K. Jain, “Angiogenesis in Cancer and Other Diseases”, *Nature*, 407(6801), 249–257, 2000. . doi:10.1038/35025220
10. W.C. Allee, E. Bowen, “Studies in Animal Aggregations: Mass Protection Against Colloidal Silver Among Goldfishes”, *Journal of Experimental Zoology*, 61(2), 185–207, 1932. doi:10.1002/jez.1400610202
11. E.A. Sarapata, L.G. de Pillis, “A Comparison and Catalog of Intrinsic Tumor Growth Models”, *Bulletin of Mathematical Biology*, 76(8), 2010–2024, 2014. doi:10.1007/s11538-014-9986-y
12. S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J.M.L. Ebos, et al., “Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth”, *PLoS Computational Biology*, 10(8), e1003800, 2014. doi:10.1371/journal.pcbi.1003800
13. A. Laird, “Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve to One Cell.”, *Br J Cancer* 19, 278–291 (1965). doi:10.1038/bjc.1965.32
14. L. von Bertalanffy, “Quantitative laws in metabolism and growth”, *The Quarterly Review of Biology*, 32(3), 217–231, 1957. doi: 10.1086/401873
15. P. Armitage, R. Doll, “The age distribution of cancer and a multi-stage theory of carcinogenesis”, *British Journal of Cancer*, 8(1), 1–12, 1954. doi: 10.1038/bjc.1954.1
16. A.C. Burton, “Rate of growth of solid tumours as a problem of diffusion”, *Growth*, 30(2), 157–176, 1966.
17. H.P. Greenspan, “Models for the growth of a solid tumor by diffusion”, *Studies in Applied Mathematics*, 51, 317–340, 1972. doi: 10.1002/sapm1972514317
18. T. Teorell, “Kinetics of distribution of substances administered to the body. I”, *Archives Internationales de Pharmacodynamie et de Thérapie*, 57, 205–225, 1937.
19. T. Teorell, “Kinetics of distribution of substances administered to the body. II”, *Archives Internationales de Pharmacodynamie et de Thérapie*, 57, 226–240, 1937.
20. R.E. Bellman, J.A. Jacquez, R. Kalaba, “Mathematical models of chemotherapy”, in: J. Neyman (ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Cambridge University Press, 57–65, 1961.
21. R.C. Rockne, J.G. Scott, “Introduction to Mathematical Oncology”, *JCO Clinical Cancer Informatics*, 3, 1–4, 2019. doi: 10.1200/CCI.19.00010
22. D. Wodarz, N. Komarova, *Dynamics of Cancer: Mathematical Foundations of Oncology*, World Scientific, 2014. doi: 10.1142/8846
23. N. Kizilova, “Mathematical modelling of biological growth and tissue engineering”, in: R. Bedzinski, M. Petryl (eds.), *Current trends in development of implantable tissue structures*, Lecture Notes of the ICB Seminar, Warsaw, 18–27, 2012.
24. N.N. Kizilova, S.A. Logvenkov, A.A. Stein, “Mathematical modeling of transport-growth processes in multiphase biological continua”, *Fluid Dynamics*, 47(1), 1–9, 2012. doi: 10.1134/S0015462812010012
25. F. Kozusko, Ž. Bajzer, “Gompertzian growth and cell population dynamics”, *Mathematical Biosciences*, 185, 153–167, 2003. doi: 10.1016/S0025-5564(03)00086-2
26. S. Xu, “Analysis of a delayed mathematical model for tumor growth”, *Nonlinear Analysis: Real World Applications*, 11, 4121–4127, 2010. doi: 10.1016/j.nonrwa.2010.04.001
27. H. Byrne, “The effect of time delays on the dynamics of avascular tumor growth”, *Mathematical Biosciences*, 144, 83–117, 1997. doi: 10.1016/S0025-5564(97)00034-0

28. M. Kamran, J.Y. Abdullah, A.S. Ahmad Satmi, M. Genisa, A. Majeed, T. Nadeem, “Mathematical modeling and analysis of tumor growth models integrating treatment therapy”, *Mathematics and Computers in Applications*, 30, 119, 2025. doi: 10.3390/mca30060119
29. H. Byrne, M.A.J. Chaplain, “Growth of necrotic tumors in the presence and absence of inhibitors”, *Mathematical Biosciences*, 135, 187–216, 1996. doi: 10.1016/0025-5564(96)00011-5
30. H.P. Greenspan, “On the growth and stability of cell cultures and solid tumors”, *Journal of Theoretical Biology*, 56, 229–242, 1976. doi: 10.1016/S0022-5193(76)80054-9
31. F.J. Solis, S.E. Delgadillo, “Discrete mathematical models of an aggressive heterogeneous tumor growth with chemotherapy treatment”, *Mathematical and Computer Modelling*, 50, 646–652, 2009. doi: 10.1016/j.mcm.2009.05.010
32. A. Qi, X. Zheng, C. Du, B. An, “A cellular automaton model of cancerous growth”, *Journal of Theoretical Biology*, 161, 1–12, 1993. doi: 10.1006/jtbi.1993.1031
33. R.A. Gatenby, R.J. Gillies, “A microenvironmental model of carcinogenesis”, *Nature Reviews Cancer*, 8(1), 56–61, 2008. doi: 10.1038/nrc2255
34. P. Magni, M. Simeoni, I. Poggesi, M. Rocchetti, G. De Nicolao, “A mathematical model to study the effects of drug administration on tumor growth dynamics”, *Mathematical Biosciences*, 200, 127–151, 2006. doi: 10.1016/j.mbs.2005.12.028
35. B. Ribba, N.H.G. Holford, P. Magni, et al., “A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis”, *CPT: Pharmacometrics & Systems Pharmacology*, 3(5), e113, 2014. doi: 10.1038/psp.2014.14
36. C. Vaghi, A. Rodallec, R. Fanciullino, et al., “Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors”, *PLoS Computational Biology*, 16(2), e1007178, 2020. doi: 10.1371/journal.pcbi.1007178
37. H.G. Müller, U. Stadtmüller, “Generalized functional linear models”, *The Annals of Statistics*, 33(2), 774–805, 2005. doi: 10.1214/009053604000001156
38. G. Lorenzo, S.R. Ahmed, D.A. Hormuth, et al., “Patient-specific, mechanistic models of tumor growth incorporating artificial intelligence and big data”, *Annual Review of Biomedical Engineering*, 26(1), 529–560, 2024. doi: 10.1146/annurev-bioeng-081623-025834
39. K. Böttger, H. Hatzikirou, A. Voss-Böhme, E.A. Cavalcanti-Adam, M.A. Herrero, A. Deutsch, “An emerging Allee effect is critical for tumor initiation and persistence”, *PLoS Computational Biology*, 11(9), e1004366, 2015. doi: 10.1371/journal.pcbi.1004366
40. J.D. Murray, *Mathematical Biology: An Introduction*, Springer, New York, 2002, pp. 1–75. doi: 10.1007/b98868
41. R. Schuster, H. Schuster, “Reconstruction models for the Ehrlich ascites tumor of the mouse”, in: O. Arino, D. Axelrod, M. Kimmel (eds.), *Mathematical Population Dynamics*, Vol. 2, Wuertz, Winnipeg, Canada, 335–348, 1995.
42. U. Foryś, A. Marciniak-Czochra, “Logistic equations in tumour growth modelling”, *International Journal of Applied Mathematics and Computer Science*, 13(3), 317–325, 2003.
43. H.H. Diebner, T. Zerjatke, M. Griehl, I. Roeder, “Metabolism is the tie: The Bertalanffy-type cancer growth model as common denominator of various modelling approaches”, *Biosystems*, 167, 1–23, 2018. doi: 10.1016/j.biosystems.2018.03.004
44. M. Kühleitner, N. Brunner, W.G. Nowak, K. Renner-Martin, K. Scheicher, “Best fitting tumor growth models of the von Bertalanffy–Pütter type”, *BMC Cancer*, 19(1), 683, 2019. doi: 10.1186/s12885-019-5911-y
45. S.S. Hassan, H.M. Al-Saedi, “Comparative study of tumor growth based on single species models”, *BIO Web of Conferences*, 97, 00118, 2024. doi: 10.1051/bioconf/20249700118
46. L. Norton, “A Gompertzian model of human breast cancer growth”, *Cancer Research*, 48, 7067–7071, 1988.
47. C. Guiot, P.G. Degiorgis, P.P. Delsanto, P. Gabriele, T.S. Deisboeck, “Does tumor growth follow a ‘universal law’?”, *Journal of Theoretical Biology*, 225, 147–151, 2003. doi: 10.1016/S0022-5193(03)00221-2
48. R.G. Cornell (ed.), *Statistical Methods for Cancer Studies*, Statistics: A Series of Textbooks and Monographs, Vol. 51, CRC Press, 1984, 496 p.
49. F. Emmert-Streib, M. Dehmer (eds.), *Statistical Diagnostics for Cancer: Analyzing High-Dimensional Data*, Quantitative and Network Biology Series, 2013, 292 p.

50. H. Ishwaran, U.B. Kogalur, E.H. Blackstone, M.S. Lauer, “Random survival forests”, *The Annals of Applied Statistics*, 2(3), 841–860, 2008. doi: 10.1214/08-AOAS169
51. M. Kamran, J.Y. Abdullah, A.S. Ahmad Satmi, M. Genisa, A. Majeed, T. Nadeem, “Mathematical modeling and analysis of tumor growth models integrating treatment therapy”, *Mathematics and Computers in Applications*, 30, 119, 2025. doi: 10.3390/mca30060119
52. Y. Suter, M. Lê, G. Glauser, et al., “The LUMIERE dataset: Longitudinal glioblastoma MRI with expert RANO evaluation”, *Scientific Data*, 9(1), Article 768, 2022. doi: 10.1038/s41597-022-01881-7

**Tiurdo Ivan** *PhD of the Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine;*

**Sediuk Anastasiia** *Student of the Faculty of Mathematics and Informatics, Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine*

**Kizilova Natalya** *Doctor of Physical and Mathematics Sciences, Professor; professor of the Department of Applied Mathematics, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine;*

## **Mathematical modeling of tumor growth dynamics for personalized therapy selection.**

Purpose of the work: to analyze current approaches to mathematical modeling of tumor growth and prediction of their dynamics using classical deterministic models and machine learning methods, as well as to determine the prospects for their use in modern mathematical oncology and personalized antitumor therapy.

Research methods: analysis and systematization of modern scientific publications on mathematical oncology; use of mathematical modeling methods for tumor growth (exponential, logistic, Gompertz and Bertalanffy models); statistical analysis of clinical data; application of machine learning methods for regression analysis and prediction of tumor growth dynamics based on longitudinal MRI data from the open LUMIERE dataset.

As a result of the study, a review and comparative analysis of classical mathematical models of tumor growth and their modifications used to describe the biological processes of proliferation and restriction of tumor tissue growth was performed. Preliminary processing and analysis of clinical and imaging data, including the volumes of various tumor components, was carried out. Individual tumor growth trajectories were modeled using regression models and ensemble machine learning methods, in particular Random Forest. It was shown that machine learning methods provide more stable and accurate predictions of complex tumor growth dynamics compared to classical models in the case of high data variability.

Conclusions: Combining classical mathematical models of tumor growth with modern machine learning methods is a promising direction for the development of mathematical oncology. This approach improves the accuracy of predicting individual tumor dynamics and creates a basis for developing personalized treatment strategies. The results obtained indicate the feasibility of further use of hybrid models in research on precision medicine and personalized antitumor therapy.

**Keywords:** *mathematical oncology, tumor growth models, Gompertz model, Bertalanffy model, machine learning, Random Forest, prediction, personalized therapy, precision medicine, clinical data.*