

Біологічний вік людини: регресійна та нейромережна моделі

А.В. Писарук¹, В.Б. Шатило¹, І.А. Антонюк-Щеглова¹, С.С. Наскалова¹,
О.В. Бондаренко¹, В.П. Чижова¹, В.В. Шатило², Л.Г. Полягушко²

¹ДУ «Інститут геронтології ім. Д. Ф. Чеботарьова НАМН України», Київ;

²НТУУ «Київський політехнічний інститут імені Ігоря Сікорського», Київ;

e-mail: elena_bondarenko@ukr.net

Мета дослідження – розробка методу визначення біологічного віку (БВ), темпу старіння людини за антропометричними та біохімічними показниками та порівняння точності за допомогою регресійного та нейромережного аналізу. У 735 практично здорових осіб віком від 20 до 79 загальноприйнятими методами визначали в плазмі крові концентрацію глюкози та інсуліну, в сироватці крові – загального холестерину, холестерину ліпопротеїдів високої, низької і дуже низької щільності, тригліцеридів, сечовини, креатиніну, трансаміназ, лужної фосфатази. Проводили антропометричні виміри, стандартний пероральний глюкозо-толерантний тест (СГТТ), розраховували індекс НОМА. Розпізнавання віку здійснювали за допомогою регресійного і нейромережного аналізу. Рівняння множинної регресії, що пов'язує вік обстежених з їх антропометричними та біохімічними лабораторними показниками, дає змогу розрахувати метаболічний вік людини з абсолютною похибкою 6,92 років. Така точність є достатньою для відбору людей із ризиком пришвидшеного старіння. Використання алгоритму нейронної мережі дає можливість визначити метаболічний вік із похибкою 4,57 років, що є достатнім для розмежування фізіологічного і пришвидшеного старіння. Використання алгоритму нейронної мережі підвищує точність визначення БВ людини, похибка його визначення зменшується на 40%.

Ключові слова: біологічний вік; антропометричні та біохімічні біомаркери; регресійна та нейромережна моделі.

ВСТУП

Загальна концепція біологічного віку (БВ) полягає в припущенні, що індивідуальна швидкість старіння може значною мірою варіювати [1–5]. БВ визначається для оцінки ризику розвитку асоційованої з віком патології та впливу профілактичних та лікувальних заходів. Розрахунок БВ здебільшого полягає у вимірі кількісних показників, іменованих біологічними маркерами старіння (БМС). Ці показники повинні досить високо корелювати з віком, бути доступними для вимірювання у клінічних умовах та мало залежати від актуальних патологічних станів. Біологічними маркерами можуть бути різні показники: анатомічні, функціональні, біохімічні, іму-

нологічні, генетичні та епігенетичні [6–9]. Нині немає загальноприйнятого набору таких маркерів.

БВ може як збігатися з хронологічним віком (ХВ), і в такому разі процес старіння оцінюється як нормальний, фізіологічний, так і значно перевищувати його, тоді старіння оцінюється як пришвидшене. Величина різниці між БВ і ХВ, яка вважається фізіологічною чи патологічною, залежить від похибки визначення БВ. Якщо ця різниця перевищує стандартну похибку, можна вважати старіння пришвидшеним. Загальноприйнятим математичним методом оцінки БВ є розрахунок рівняння множинної регресії (MLR), що пов'язує ХВ та низку кількісних показників [10–13]. Для визначення коефіцієнтів цього

рівняння використовують дані обстеження великої кількості практично здорових людей різного віку. Щоб визначити БВ, обстежуваному проводять низку досліджень та за отриманим рівнянням розраховують вік. При значному перевищенні розрахункового віку людини над ХВ робиться висновок про пришвидшене старіння. Для цього зазвичай використовують численні інструментальні та лабораторні методи дослідження. Останніми роками для визначення БВ дедалі більше використовують методи штучного інтелекту [13–17]. Це насамперед математичні методи, основані на використанні штучних нейронних мереж для аналізу даних. Революція у використанні нейронних мереж – Deep Neural Network (DNN) швидко поширюється у медицині. Такі системи виникли й у вимірюванні темпу старіння. Так, система AGE.AI (<http://aging.ai>) дає змогу досить точно оцінити БВ людини за допомогою даних, отриманих при звичайних клінічних лабораторних дослідженнях.

Можливість навчання – одна з головних переваг нейронних мереж перед традиційними алгоритмами, при цьому знаходяться коефіцієнти зв'язків між нейронами. У процесі навчання нейронна мережа здатна виявляти складні нелінійні залежності між вхідними та вихідними даними. ХВ людей визначається як змінна на виході нейромережі. В процесі навчання мережа шукає залежності між вхідними та вихідними даними, змінюючи вагові коефіцієнти зв'язків між нейронами. Нейронну мережу використовують для визначення віку людей за даними їх БМС. Якщо вона розпізнає вік людини, який істотно перевищує ХВ, то можна зробити висновок про пришвидшене старіння. Тому вік людини, розрахований нейронною мережею, логічно вважати його БВ. Застосування нейронних мереж для оцінки БВ показало їх високу ефективність за всіма критеріями, зокрема відсутність систематичної похибки, властивої методам на основі MLR [5, 9].

Для використання методу оцінки БВ у клінічній практиці він повинен включати

прості тести, доступні у звичайних медичних закладах. І тому добре підходять дані антропометричних та лабораторних досліджень пацієнтів. Раніше ми зробили таку спробу на невеликій групі людей віком від 40 до 80 років [18]. Отримана в результаті формула розрахунку БВ давала велику помилку. У цій роботі використовували результати визначення більшої кількості лабораторних показників у великій групі людей більш широкого вікового діапазону (20–80 років).

Мета нашої роботи полягала в розробці методу визначення БВ та темпу старіння людини за антропометричними та біохімічними лабораторними показниками, а також у порівнянні точності за допомогою регресійного та нейромережного методів аналізу.

МЕТОДИКА

Дослідження виконано відповідно до вимог дотримання етичних норм і принципів Гельсінської декларації. Кожен обстежений добровільно підписав форму згоди на участь у дослідженні. Відповідно до протоколу дослідження були відібрані люди віком від 20 до 79 років (735 осіб) без клінічно значущої патології серцево-судинної, дихальної, ендокринної, центральної нервової системи, без хронічних захворювань печінки і нирок, патології кровотворної системи.

Антропометричні вимірювання включали визначення маси тіла, зросту, індексу маси тіла (ІМТ), обводу талії та стегон (ОТ, ОС відповідно), співвідношення ОТ/ОС. Масу тіла визначали стаціонарною вагою з точністю до 0,1 кг в осіб у легкому одязі без взуття. Зріст вимірювали з точністю до 0,5 см в осіб без взуття. ІМТ вираховували як відношення маси тіла до зросту. Сантиметровою стрічкою вимірювали ОТ у стоячому положенні без одягу на талії після звичайного видиху – посередині між задньою частиною бічної реберної дуги та гребенем клубової кістки (стандарт ВООЗ). ОС вимірювали на рівні великого вертлюга стегнової кістки. Також

розраховували співвідношення ОТ/ОС. Значення його для чоловіків >1,0 і жінок >0,85 свідчить про абдомінальний тип ожиріння.

Проводили забір крові натще та в сироватці крові визначали вміст загального холестерину (ЗХ), холестерину ліпопротеїдів низької щільності (ХС ЛПНЩ), холестерину ліпопротеїдів дуже низької щільності (ХС ЛПДНЩ) і холестерину ліпопротеїдів високої щільності (ХС ЛПВЩ), тригліцеридів, активність трансаміназ, вміст білірубину, креатиніну і сечовини на автоматичному біохімічному аналізаторі «Autolab» фірми «Boehringer Mannheim, Іспанія». Також проводили стандартний пероральний глюкозо-толерантний тест (СГТТ). Концентрацію глюкози в плазмі крові визначали глюкозооксидазним методом на напівавтоматичному біохімічному аналізаторі BTS-330, використовуючи реагенти «Глюкоза» набору «Bio LATEST Lachema Diagnostica» (Іспанія). Вміст інсуліну в плазмі крові визначали імуноферментним методом, використовуючи набори DRG Insulin ELISA («DRG Instruments GmbH», Німеччина). Для оцінки інсулінорезистентності розраховували індекс НОМА-ІР (Homeostasis Model Assessment for Insulin Resistance) за формулою:

НОМА-ІР = глюкоза натще (ммоль/л) × інсулін натще (ммоль/л) / 22,5.

Для аналізу отриманих результатів використовували методи варіаційної статистики за допомогою програми Statistica 7.0 (StatSoft, США). Середні значення показників у різних вікових групах порівнювали за допомогою однофакторного ANOVA за фактором «Вік». Вірогідними вважалися відмінності при $P < 0,05$. Для розробки моделей БВ застосовували методи покрокової множинної регресії та нейромережного аналізу за допомогою програми Statistica 7.0.

РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

У табл. 1 наведено середні значення показників у різних вікових групах.

Використання покрокової множинної

регресії дало змогу відібрати найбільш інформативні показники і отримати рівняння, що пов'язує вік обстежених людей з низкою показників (табл. 2).

Для написання формули MLR використовуємо колонку b табл. 2, яка містить коефіцієнти регресії для кожного показника:

$$Y = 1,137 X_1 + 0,293 X_2 + 0,877 X_3 + 3,632 X_4 + 1,509 X_5 + 0,322 X_6 - 32,84$$

де:

Y – БВ, роки;

X_1 – ОТ/зріст²;

X_2 – інсулін натще, мкОД/мл;

X_3 – глюкоза через 2 год стандартного перорального глюкозо-толерантного тесту, ммоль/л;

X_4 – тригліцериди, ммоль/л;

X_5 – сечовина, ммоль/л;

X_6 – лужна фосфатаза, од./л.

Коефіцієнт кореляції предикторів із залежним показником (вік) $r = 0,848$ свідчить про високий кореляційний зв'язок між ними. Коефіцієнт детермінації моделі $r^2 = 0,721$ показує, що 72,1% дисперсії віку можна пояснити за допомогою предикторів, що увійшли в модель. Критерій Фішера $F = 317,6$ ($P < 0,0001$) свідчить про високу значущість моделі, за допомогою якої можна розраховувати БВ з достатньою точністю. Для оцінки вірогідності входження предиктора в модель розраховують Р (остання колонка табл. 2). Знак перед предиктором характеризує спрямованість його кореляційного зв'язку із залежним показником, тобто з БВ.

На рис. 1 наведено графік співвідношення БВ та ХВ. Видно, що розкид точок навколо лінії регресії невеликий і коефіцієнт множинної кореляції високий ($r = 0,848$; $P < 0,0001$). Середня абсолютна похибка розрахунку віку становила $6,92 \pm 5,46$ років.

Модель нейронної мережі БВ (DNN). За допомогою нейронної мережі, що моделюється у програмі Statistica 7.0, було вирішено завдання розпізнавання віку обстежених людей за біохімічними показниками крові. Для навчання нейронної мережі на її вхід подава-

Таблиця 1. Середні значення показників та їх стандартне відхилення (M±SD) у різних вікових групах

Показники	20–39 років (n = 253)	40–59 років (n = 252)	60–79 років (n = 230)
Обвід талії, см	72,70±6,74	91,71±8,79	95,67±10,65
Обвід стегон, см	88,9±6,52	105,1±8,06	106,68±7,85
Обвід талії / обвід стегон	0,818±0,056	0,872±0,051	0,896±0,072
Маса тіла, кг	65,6±7,74	82,5±10,78	81,8±11,70
Зріст, м	1,70±0,069	1,70±0,076	1,65±0,086
Індекс маси тіла, кг/м ²	22,5±1,53	28,7±3,41	30,2±4,19
Глюкоза натще, ммоль/л	5,39±0,37	5,80±0,57	5,90±0,74
Глюкоза через 2 год стандартного перорального глюкозо-толерантного тесту, ммоль/л	5,76±0,78	6,41±1,17	7,36±1,89
Інсулін натще, мкОД/мл	6,73±1,53	10,94±4,59	12,34±5,37
Індекс НОМА	1,62±0,39	2,84±1,32	3,28±1,53
Загальний холестерин, ммоль/л	4,74±0,56	5,57±0,90	5,81±1,05
Тригліцериди, ммоль/л	0,76±0,18	1,53±0,54	1,74±0,63
Холестерин ліпопротеїдів високої щільності, ммоль/л	1,44±0,097	1,34±0,15	1,34±0,24
Холестерин ліпопротеїдів низької щільності, ммоль/л	2,81±0,50	3,47±0,80	3,59±0,92
Холестерин ліпопротеїдів дуже низької щільності, ммоль/л	0,38±0,10	0,76±0,27	0,88±0,33
Креатинін, мкмоль/л	78,3±8,98	80,5±10,11	83,4±16,04
Сечовина, ммоль/л	5,09±0,80	5,52±1,00	5,97±1,47
Лужна фосфатаза, од./л	60,8±9,35	79,9±15,5	91,1±12,5

Примітка: вплив фактора «Вік» є достовірним для всіх показників (ANOVA, P < 0,0001).

ли результати обстеження людей та їх ХВ, а виходом був прогнозований (біологічний) вік. У процесі навчання мережа шукає залежності

між вхідними значеннями та ХВ, змінюючи вагові коефіцієнти зв'язків між нейронами. Найкращі результати показала нейронна ме-

Таблиця 2. Результати розрахунку рівняння MLR у програмі Statistica 7.0 (n = 735)

Показники	b*	SE of b*	b	SE of b	t (200)	P
Константа			-32,84	2,380	-13,81	0,0001
Обвід талії/зріст, см/м ²	0,382	0,027	1,137	0,081	14,08	0,0001
Інсулін, мкОД/мл	0,081	0,023	0,293	0,082	3,56	0,0004
Глюкоза через 2 год стандартного перорального глюкозо-толерантного тесту, ммоль/л	0,076	0,022	0,877	0,256	3,42	0,0007
Тригліцериди, ммоль/л	0,137	0,026	3,632	0,683	5,32	0,0001
Сечовина, ммоль/л	0,103	0,021	1,509	0,302	4,99	0,0001
Лужна фосфатаза, од./л	0,336	0,025	0,322	0,024	13,25	0,0001

Примітка: b – коефіцієнт регресії; b* – стандартизований коефіцієнт регресії; SE of b (b*) – стандартизована похибка коефіцієнта.

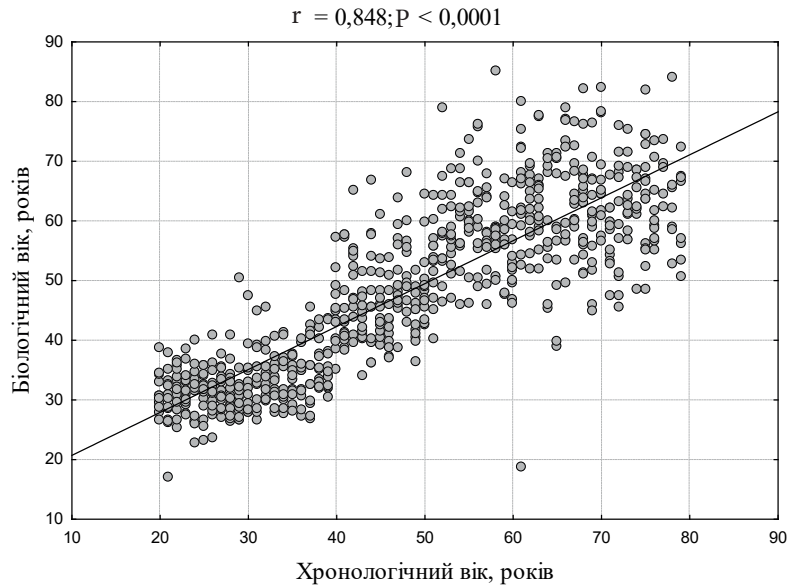


Рис. 1. Співвідношення між біологічним і хронологічним віком людей (MLR-модель)

режа MLP з 2 внутрішніми шарами з 15 та 12 нейронами (рис. 2).

Інформативність відібраних показників, використаних для розпізнавання віку людини, представлено в табл. 3. Найвищий ранг мають такі показники: вміст лужної фосфатази, ОТ та тригліцеридів.

На рис. 3 наведено графік співвідношення БВ та ХВ (DNN-модель). Видно, що розкид

точок навколо лінії регресії дуже невеликий і коефіцієнт множинної кореляції високий ($r = 0,940; P < 0,0001$). Середня абсолютна похибка розрахунку віку становила 4,57 років.

Порівняння точності оцінки БВ за допомогою регресійної (похибка 6,92 років) та нейромережної (похибка 4,57 років) моделей виявило істотну перевагу нейромережі. Однак для використання навченої нейромережі для

Профіль нейронної мережі MLP

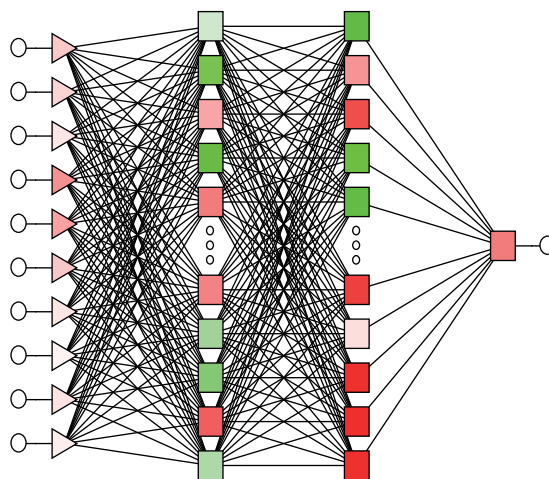


Рис. 2. Структура нейронної мережі для обчислення біологічного віку: вхідний шар нейронів (показники), два внутрішніх шари та вихідний нейрон (вік)

Таблиця 3. Інформативність показників розпізнавання віку

Показники	Інформативність	Ранг
Лужна фосфатаза	1,45	1
Обвід талії/зріст ²	1,43	2
Тригліцериди	1,25	3
Сечовина	1,22	4
Індекс маси тіла	1,21	5
Глюкоза натще	1,17	6
Глюкоза через 2 год стандартного перорального глюкозо-толерантного тесту	1,13	7
Інсулін натще	1,12	8
Загальний холестерин	1,10	9

розрахунку БВ конкретної людини потрібна спеціальна програма, тоді як застосування регресійної формули загальнодоступне.

ВИСНОВКИ

1. Рівняння множинної регресії, що пов’язує вік обстежених з їх антропометричними та біохімічними лабораторними показниками, дає змогу розрахувати вік людини з абсолютною похибкою 6,92 років. Така точність є достатньою для відбору людей із ризиком

пришвидшеного старіння.

2. Використання алгоритму нейронної мережі дає можливість визначити БВ із похибкою 4,57 років, що свідчить про переваги цього методу порівняно з регресійним.

3. Розроблений метод визначення БВ може знайти застосування для оцінки ризику розвитку патології, ефективності лікувально-профілактичних та реабілітаційних заходів. *Роботу виконано в межах планової наукової тематики ДУ «Інститут геронтології ім. Д.Ф. Чеботарьова НАМН України» «Розроби-*

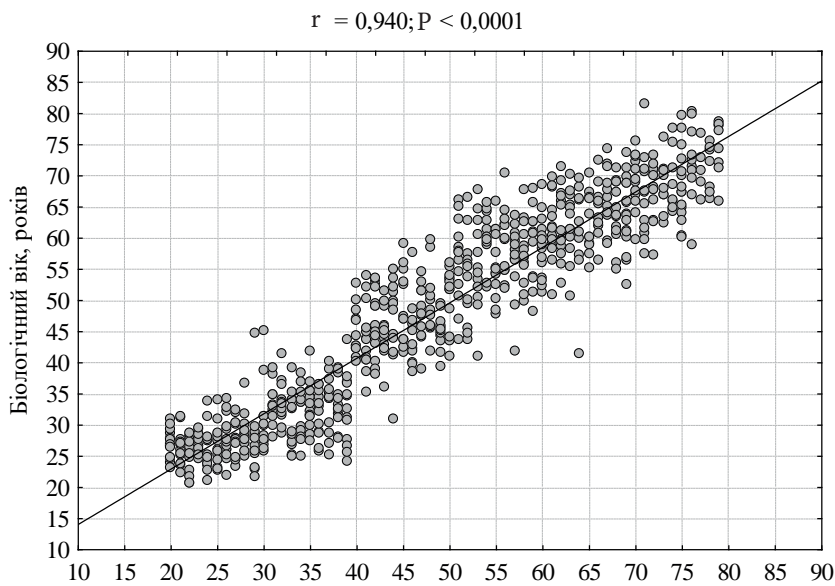


Рис. 3. Співвідношення між біологічним і хронологічним віком людей (DNN-модель)

ти технологію діагностики темпу і профілю старіння людини з використанням нейромережових алгоритмів» (державний реєстраційний № 0119U103980). Установою, що фінансує дослідження, є НАМН України.

The authors of this study confirm that the research and publication of the results were not associated with any conflicts regarding commercial or financial relations, relations with organizations and/or individuals who may have been related to the study, and interrelations of co-authors of the article.

**A.V. Pisaruk¹, V.B. Shatilo¹,
I.A. Antoniuk-Shcheglova¹, S.S. Naskalova¹,
O.V. Bondarenko¹, V.P. Chyzhova¹, V.V. Shatilo²,
L.G. Polyagushko²**

HUMAN BIOLOGICAL AGE: REGRESSION AND NEURAL NETWORK MODELS

¹*D.F. Chebotarev Institute of Gerontology NAMS of Ukraine, Kyiv;*

²*Ihor Sikorsky Kyiv Polytechnic Institute, Kyiv;
e-mail: elena_bondarenko@ukr.net*

The purpose of this study is to develop methods for determining the biological age (BA) and the pace of human aging based on anthropometric and biochemical laboratory biomarkers, comparing the accuracy of BA determination using regression and neural network analysis. In 735 practically healthy people aged from 20 to 79, we determined the blood plasma concentrations of glucose and insulin and the blood serum concentrations of total cholesterol, cholesterol of high, low and very-low density lipoproteins, triglycerides, urea, creatinine, transaminases, and alkaline phosphatase. Also, we conducted anthropometric measurements and a standard oral glucose tolerance test and calculated the HOMA index. Age recognition was carried out using regression and neural network analysis. The multiple regression equation, which connects the examinees age with their biochemical laboratory parameters, allows to calculate the metabolic age of a person with an absolute error of 6.92 years. This accuracy is sufficient to reveal people at risk of accelerated aging. The use of a neural network algorithm with deep learning allows to determine the metabolic age with an error of 4.57 years, which is sufficient to distinguish between physiological and accelerated aging. The use of a neural network algorithm with deep learning increases the accuracy of determining a person's metabolic age, the error of its determination is reduced by 40%.

Key words: biological age; anthropometric and biochemical biomarkers; regression and neural network models.

REFERENCES

1. Korkushko OV, Shatilo VB. Accelerated aging and its prevention. *Bukovyn Med Bull.* 2009;13(4):153-8. [Ukrainian].
2. Cevenin EI, Invidia L, Lescai F, Salvioli S, Tieri P, Castellani G, Franceschi C. Human models of aging and longevity. *Exp Opin Biol Ther.* 2008;8(9): 1393-405.
3. Mitnitski AB, Graham JE, Mogilner AJ, Rockwood K. Frailty, fitness and late-life mortality in relation to chronological and biological age. *BMC Geriatr.* 2002 Feb 27;2:1.
4. Morgan E. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol A Biol Sci Med Sci.* 2013;68(6):667-74.
5. Salthouse T. Aging and measures of processing speed. *Biol Psychol.* 2000;54(1):35-54.
6. Bürkle A, Moreno-Villanueva M, Bernhard J, Blasco M, Zondag G, Hoeijmakers JH, et al. MARK-AGE biomarkers of ageing. *Mech Ageing Dev.* 2015 Nov;151:2-12.
7. Cardoso AL, Fernandes A, Aguilar-Pimentel JA, de Angelis MH, Guedes JR, Brito MA, et al. Towards frailty biomarkers: Candidates from genes and pathways regulated in aging and age-related diseases. *Ageing Res Rev.* 2018 Nov;47:214-77.
8. Moreno-Villanueva M, Capri M, Breusing N, Siepelmeyer A, Sevini F, Ghezzi A, de Craen AJ, Hervonen A, Hurme M, Schön C, Grune T, Franceschi C, Bürkle A. MARK-AGE standard operating procedures (SOPs): A successful effort. *Mech Ageing Dev.* 2015 Nov;151:18-25.
9. Xia X, Chen W, McDermott J, Han JJ. Molecular and phenotypic biomarkers of aging. *F1000Res.* 2017 Jun 9;6:860.
10. Aykroyd RG, Lucy D, Pollard AM, Solheim T. Technical note: regression analysis in adult age estimation. *Am J Phys Antropol.* 1997;104:259-65.
11. Jia L, Zhang W, Chen X. Common methods of biological age estimation. *Clin Interv Aging.* 2017;12:759-72.
12. Kröll J, Saxtrup O. On the use of regression analysis for the estimation of human biological age. *Biogerontology.* 2000;1:363-68.
13. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. *Mol Pharm.* 2016;13(5):1445-54.
14. Caballero FF, Soulis G, Engchuan W, Sánchez-Niubó A, Arndt H, Ayuso-Mateos JL, Haro JM, Chatterji S, Panagiotakos DB. Advanced analytical methodologies for measuring healthy ageing and its determinants, using factor analysis and machine learning techniques: the ATHLOS project. *Sci Rep.* 2017 Mar 10;7:43955.
15. Cole JH, Franke K, Cherbuin N. Quantification of the biological age of the brain using neuroimaging. In: *Biomarkers of Human Aging.* Moskalev A (ed.). Switzerland: Springer; 2019. p. 293-328.
16. Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, Ostrovskiy A, Cantor C, Vijg J, Zhavo-

- ronkov A. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. *Aging (Albany NY)*. 2016 May;8(5):1021-33.
17. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. *J Machine Learn Res*. 2014;15(56):1929-58.
18. Korkushko OV, Pizaruk AV, Chyzhova VP. Estimation of human metabolic age using regression and neural network. *Zaporozh Med J*. 2021;23(1):60-64.

*Матеріал надійшов до
редакції 01.02.23*