

РОЗДІЛ II. ІНФОРМАЦІЙНО-КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

DOI: 10.25140/2411-5363-2024-1(35)-121-127

УДК 616-085:311.218

Володимир Єременко¹, Олена Монченко², Тарас Монченко³, Валентина Кучеренко⁴

¹доктор технічних наук, завідувач кафедри інформаційно-вимірювальних технологій
Національний технічний університет України

«Київський політехнічний інститут ім. Ігоря Сікорського» (Київ, Україна)

E-mail: nau_307@ukr.net. **ORCID:** <https://orcid.org/0000-0002-4330-7518>. Scopus Author ID: 56736964700

²кандидат технічних наук, доцент кафедри біокібернетики та аерокосмічної медицини
Національний авіаційний університет (Київ, Україна)

E-mail: monchenko_olena@ukr.net. **ORCID:** <https://orcid.org/0000-0002-8248-5704>.

ResearcherID: [AAN-1910-2020](https://orcid.org/0000-0002-8248-5704). **ID Scopus:** 56422110000

³здобувач вищої освіти кафедри акустичних та мультимедійних електронних систем
Національний технічний університет України

«Київський політехнічний інститут ім. Ігоря Сікорського» (Київ, Україна)

E-mail: mtm082220-ames27@iit.kpi.ua. **ORCID:** <https://orcid.org/0009-0002-3294-7023>

⁴кандидат технічних наук, доцент кафедри біокібернетики та аерокосмічної медицини
Національний авіаційний університет (Київ, Україна)

E-mail: valentyna.kucherenko@npp.nau.edu.ua. **ORCID:** <https://orcid.org/0000-0003-1009-8729>

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ОПРАЦЮВАННЯ МЕДИЧНИХ ПОКАЗНИКІВ

Розроблено інформаційну технологію обробки медичних показників шляхом застосування гістограмного аналізу. Проведено аналіз гістограмних досліджень при обробці потоків даних. Розроблено алгоритм роботи, який складається з двох етапів, проаналізовано отримані результати. На першому етапі побудовано гістограми для всіх медичних показників. На другому етапі проаналізовано ймовірність зміни кожного показника, що дало можливість перейти від конкретного пацієнта до ймовірності в досліджуваній групі та проаналізувати: збільшилась чи зменшилась ймовірність впливу певним методом лікування. Використана інформаційна технологія застосування гістограмних чисел для обрахунку медичних показників і визначення частки пацієнтів, для яких лікування виявилось ефективним. Окреслені основні етапи розробки, що дають змогу здійснювати гістограмний аналіз і приймати рішення щодо подальших досліджень.

Ключові слова: гістограмний аналіз; інформаційна технологія; медичні показники.

Рис.: 3. Табл.: 1. Бібл.: 7.

Актуальність теми дослідження. Гістограмний аналіз є одним із сучасних методів графічного аналізу даних. З історичної перспективи метод цей заснований на тезисі, що варіація показників результатів вимірювань є інтегральною. Цілком зрозуміло, що виходячи виключно з цього твердження не можливо отримати інформацію про характер помилок, таким чином не можливо зрозуміти, чи виникли помилки дослідженні статистично, чи є наслідком якоїсь проблеми, адже відповіді на це інтегральні параметри дати не можуть.

Гістограмний аналіз є важливим інструментом у статистичному аналізі для дослідження вибіркового закону розподілу. Цей метод використовує гістограму для визначення частоти або кількості спостережень у різних інтервалах або «кошиках» даних.

Гістограмний аналіз застосовується в багатьох сферах. До прикладу, дослідження варіації даних: гістограми допомагають побачити як розподіл даних, чи є він нормальним, або відповідає одному з теоретичних законів розподілу тощо. Виявлення аномалій: завдяки гістограмам можна виявляти викиди або аномальні значення в наборах даних. Відсутність або наявність певних фрагментів, вимірювання центральних моментів, порівняння розподілів, попередній аналіз перед статистичними тестами, оптимізація бізнес-процесів, аналіз лабораторних показників, дослідження захворювань, фармакологічні дослідження, оцінка ризиків, епідеміологічні дослідження, класифікація даних тощо.

Загалом гістограмний аналіз також є потужним інструментом для аналізу медичних даних, дозволяючи відкривати закономірності, виявляти аномалії та забезпечувати підґрунтя для подальших досліджень.

Постановка проблеми. Аналіз існуючих алгоритмів подання та опрацювання медичної інформації показує, що вони не повною мірою задовольняють вимоги до розв'язання задач, що вимагають складних логічних висновків з урахуванням неповноти та суперечливості вхідних даних. Для вирішення вказаної проблеми необхідно шукати нові технології опрацювання медичних даних.

Для постановки правильного діагнозу та призначенні ефективного лікування лікарям доводиться знімати велику кількість різних показників: антропогенних, по крові тощо.

На підставі цього, необхідно розробити інформаційну технологію, яка б допомагала практикуючим лікарям підвищити якість прийнятих рішень щодо вибору схеми лікування з урахуванням індивідуальних особливостей хворого. Ефективна інформаційна стратегія запобігає появі небажаних наслідків лікування (а саме алергічної реакції на фармакологічні засоби) та сприяє зменшенню собівартості послуг, а також покращує їхню ефективність, завдяки чому лікувальні установи можуть значно покращити рівень своєї роботи.

У статті [1] показані нефармакологічні стратегії, спрямовані на зміни у способі життя, що можуть позитивно впливати на артеріальний тиск (АТ). Доведено за рахунок визначення відстані Махаланобіса, що для лікування артеріальної гіпертензії (АГ) та ожиріння ефективно показав себе новий метод лікування DASH. Таким чином, DASH може бути рекомендованим методом для зниження АТ у дорослих з попереднім етапом гіпертензії. Проте ці технології фактично визначають інтегральні параметри, які не дають можливості окремо аналізувати впливи на медичні показники.

Постала задача детального аналізу кожного отриманого показника для виявлення впливу лікування на кожен показник з подальшим аналізом та прогнозом щодо перебігу лікування.

У статті запропоновано використання гістограмного аналізу кожного показника по досліджуваних групах: група здорових пацієнтів (M1), група хворих пацієнтів, які приймали лікування DASH (M2), та група хворих пацієнтів, які лікувались стандартним методом (M3).

Аналіз останніх досліджень і публікацій. У статті [1] запропоновано інформаційну технологію, що полягає у статистичній обробці великої кількості різнопланових даних, що ґрунтується на використанні відстані Махаланобіса. Такий підхід дав загальний результат і підтвердив ефективність нових методів лікування артеріальної гіпертензії та надлишкової ваги.

У роботі [2] розроблено інформаційну технологію обробки персоналізованої медичної інформації для прийняття рішень і як засіб її реалізації та запропоновано систему підтримки прийняття лікарських рішень, розроблено її архітектуру, проаналізовано отримані результати.

В роботі [3] запропонований метод гістограмного аналізу, який підходить для інкрементального навчання, а також для задач, де потоки даних неперервні, як у більшості сучасних додатків для інтелектуального аналізу даних.

У статті [4] використаний гістограмний аналіз при обробці гістологічних і цитологічних зображень.

Проведений аналіз досліджень і публікацій показав, що гістограмний аналіз дає гарні результати при обробці потоків даних. У медичних дослідженнях інформаційна технологія, що ґрунтується на застосуванні гістограмного аналізу медичних показників, практично не використовується.

У роботі пропонується інформаційна технологія гістограмного аналізу ендокринологічних показників при лікуванні артеріальної гіпертензії та надлишкової ваги.

Виділення недосліджених частин загальної проблеми. Для дослідження розподілу ваги в дослідній групі та отримання інформації про характеристики цього розподілу, такі як середнє значення, медіана, дисперсія та інші параметри необхідно провести додаткові дії. Це може допомогти визначити, наскільки поширена проблема надлишкової ваги в

певній групі та розробити ефективність стратегії лікування. Зокрема, гістограмний аналіз може допомогти визначити частку пацієнтів, у яких показники в нормі, або виходять за межі норми, що дає можливість використати стандартні методики статистичних досліджень і може бути корисним для розробки ефективних методів лікування.

Мета статті – розробити інформаційну технологію аналізу ймовірності зміни значень медичних показників в процесі обраного методу лікування.

Виклад основного матеріалу. Розглянемо окремих випадок прийняття рішення за результатами вимірювання слабоінформативних параметрів. Припустимо, що задані еталонні вибірки вимірювань одного слабо інформативного параметра, яким є певний медичний показник: $x_1(k)$ і $x_2(k)$, $k=1,2,\dots,n$. Визначимо за об'єднаною вибіркою вимірювань мінімальне $x_{\min}=\xi_1$ значення та максимальне – $x_{\max}=\xi_n$. Виберемо число гістограмних інтервалів L та обчислимо розмір елементарного інтервалу дискретизації $\Delta x = \frac{(\xi_n - \xi_1)}{L}$. За вибірками вимірювань визначимо гістограмні числа еталонних вимірювань об'єктів O_1 і O_2 : $k_{11}, k_{12}, \dots, k_{1l}$ і $k_{21}, k_{22}, \dots, k_{2l}$.

Гістограмні числа k_1, k_2, \dots, k_l є випадковими величинами. За допомогою гістограмних чисел визначаємо частку пацієнтів, для яких лікування виявилось ефективним. Це можна також трактувати як ймовірність ефективності лікування. Якщо відомий закон розподілу результатів вимірювання $W(x)$, то можна визначити ймовірність попадання в i -тий інтервал гістограми одного вимірювання

$$p_i = \int_{x_{oi} - \frac{\Delta x}{2}}^{x_{oi} + \frac{\Delta x}{2}} W(x) dt \approx W(x_{oi}) \Delta x$$

та записати поліноміальний закон розподілу гістограмних чисел

$$P_n(k_1, k_2, \dots, k_l) = \frac{n! \prod_{i=1}^l p_i^{k_i}}{\prod_{i=1}^l (k_i)!} \quad (1)$$

Гістограмні числа є залежними випадковими величинами, тому що їх сума $\sum_{i=1}^l k_i = n$. Кожне з чисел k_i підпорядковується біноміальному закону розподілу ймовірностей. Оскільки $M[k_i] = np_i$, $D[R_i] = np_i(1-p_i)$, $M[k_i k_j] = n^2 p_i p_j - np_i p_j$, то коефіцієнт кореляції r_{ij} , дорівнює

$$r_{ij} = - \frac{\sqrt{p_i p_j}}{\sqrt{(1-p_i)(1-p_j)}}.$$

Таким чином, якщо об'єкти розпізнавання O_1 і O_2 , описуються багатовимірними законами розподілу $W(x^k/O_1)$ та $W(x^k/O_2)$, то їх гістограми описуються поліноміальними законами розподілу $P(k^k/O_1)$ та $P(k^k/O_2)$ з параметрами p_{1i} та p_{2i} , $i = 1, 2, \dots, L$. Гістограмні числа еталонних вимірювань об'єктів O_1 і O_2 є оцінками ймовірностей p_{1i} та p_{2i}

$$p_{1i}^* = \frac{k_{1i}}{n} = p_{1i}^*, \quad p_{2i}^* = \frac{k_{2i}}{n} = p_{2i}^*. \quad (2)$$

За еталонні вимірювання прийемо показники, зняті зі здорової групи досліджуваних $M1$ p_{11}^* . Гістограмні числа показують ймовірність значення показника у групі досліджуваних пацієнтів. Досліджувана група, що приймала лікування 1, має гістограмні показники: до лікування p_{21}^* , після лікування p_{22}^* . Досліджувана група, що приймала лікування 2, має гістограмні показники: до лікування p_{31}^* , після лікування p_{32}^* .

Різниця гістограмних чисел $\Delta_2 = p_{22}^* - p_{21}^*$, $\Delta_3 = p_{32}^* - p_{31}^*$.

Матеріали та методи досліджень. Для виконання запропонованої інформаційної технології в роботі був запропонований алгоритм, наведений на рис. 1.

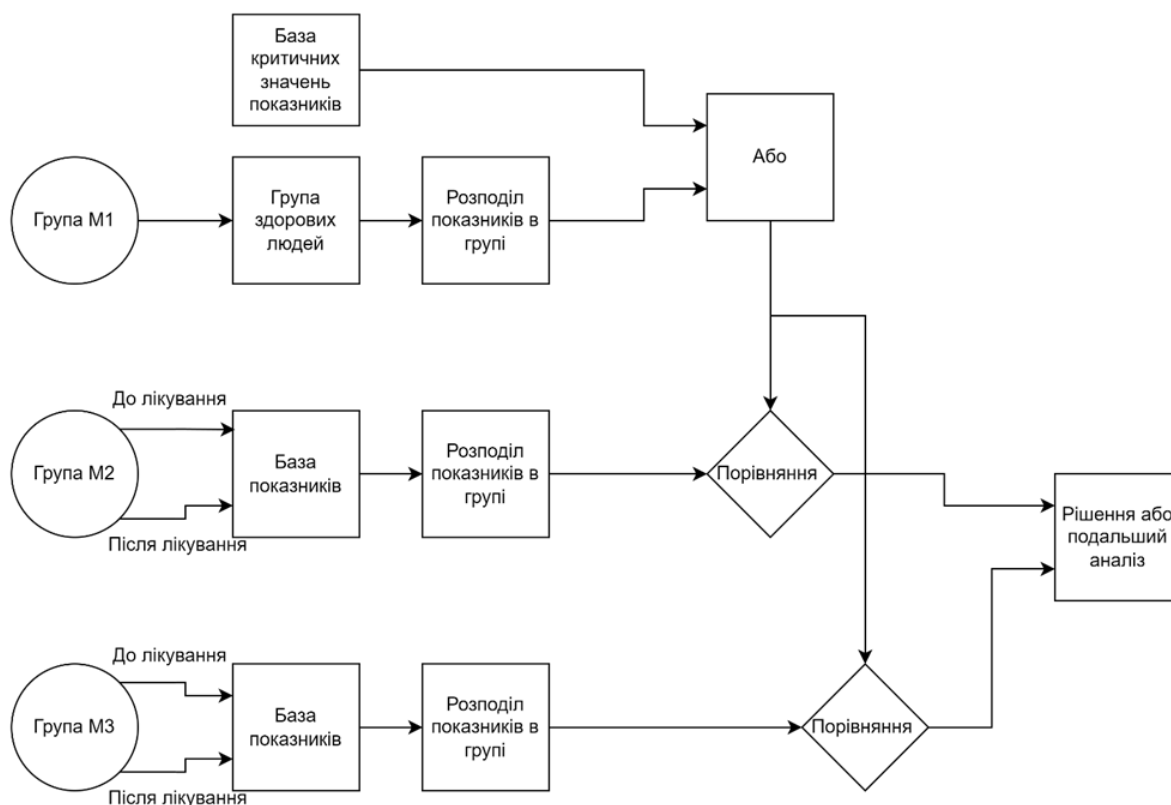


Рис. 1. Алгоритм роботи інформаційної технології обрахунку медичних показників гістограмним методом

У процесі проведення роботи отримано дані до та після лікування надлишкової ваги трьох груп пацієнтів. На основі цих матеріалів було побудовано гістограми розподілу даних. З 32 показників отримано сумарно близько 150 розподілів, які допомагають наочно побачити зміни в показниках пацієнтів. Першим етапом досліджень була побудова гістограм по всіх показниках з нормуванням за формулами (1), (2) для представлення загальної картини до та після лікування двома типами лікування 1 і 2. Фрагмент побудованих гістограм наведений на рис.2 (вісь X – значення показника; вісь Y – кількість пацієнтів):

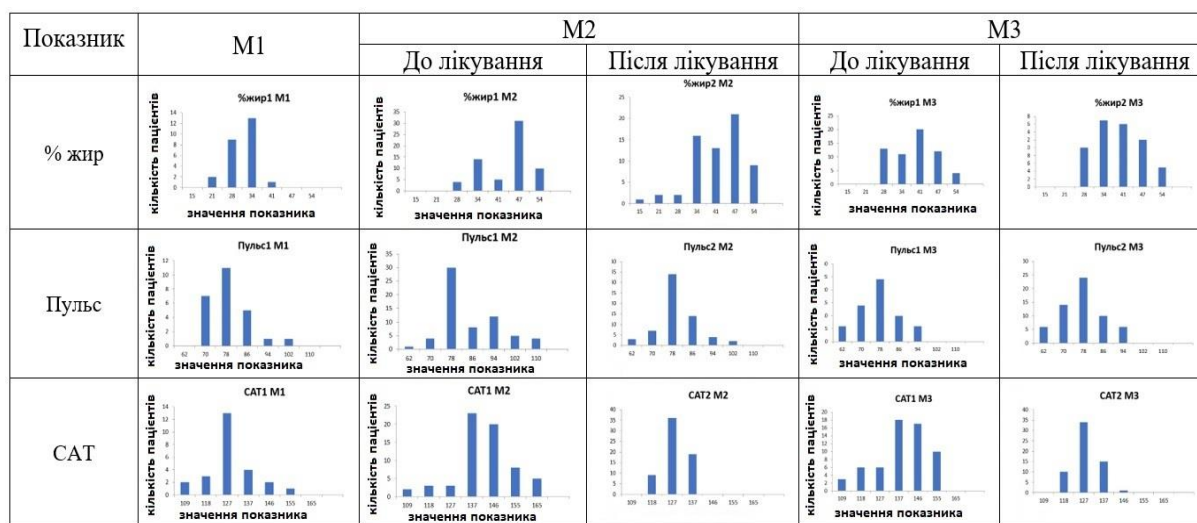


Рис. 2. Фрагмент нормованих гістограм

Другим етапом був проведений розрахунок ймовірності щодо покращення або погіршення рівня показника. Особливістю обрахунку є те, що нема прив'язки до конкретного пацієнта. Підрахунок відбувався за формулами (1), (2), що дало можливість перейти від конкретного пацієнта до ймовірності в досліджуваній групі та проаналізувати «тонку» структуру: збільшилась чи зменшилась ймовірність впливу певним методом лікування.

Результати досліджень. До прикладу розглянемо структуру зміни ймовірності в показнику «Пульс» до та після лікування методикою 1 (група M2) (рис. 3):



Рис. 3. Зміна ймовірності в показнику «Пульс» до та після лікування методикою 1

Як видно з рис. 3, ймовірність появи пацієнтів зі значеннями пульсу 110 після лікування дорівнює нулю. Також ймовірність появи пацієнтів зі значеннями пульсу 102 до лікування була 0,06, після лікування ймовірність знизилась до 0,03. Ймовірність значення пульсу 94 до лікування 0,16, після лікування 0,03. Отримані результати показують те, що лікування за методикою 1 для нормалізації пульсу є підходящим.

Також запропонована в роботі інформаційна технологія гістограмного аналізу медичних показників виявила показники, які показали суттєве погіршення після обох методик лікування, а значить необхідно звернути увагу на ці показники і провести додаткові дослідження.

Третій етап досліджень був більш інтегральним і полягав у порівнянні отриманих після лікування обома методиками гістограм з пороговими значеннями показників.

Результати розрахунку ймовірності за всіма показниками наведені в таблиці 1.

Таблиця 1 – Результати розрахунку ймовірності за показниками

Показники	p ₁₁ *	M2		Δ ₂	M3		Δ ₃
		p ₂₁ *	p ₂₂ *		p ₃₁ *	p ₃₂ *	
Вага	0,6	1,45	1,45	0	1,45	1,45	0
ІМТ	0,44	0,52	0,52	0	0,52	0,52	0
% жиру	1,36	0,22	0,64	0,042	0,17	0,26	0,09
Пульс	0,44	0,47	0,53	0,06	0,375	0,375	0
САТ	5,08	0,05	0,56	0,51	0,09	0,53	0,44
ДАТ	0,72	0,45	0,78	0,33	0,4	0,79	0,39
Креат сечі	0,64	0,52	1	0,48	0,48	0,94	0,46
Гліков Нб	0,96	0,22	0,125	-0,095	0,186	0,078	-0,11
Сеч к-та	0,6	0,26	0,16	-0,1	0,25	0,14	-0,11
Глюкоза	0,76	0,83	0,84	0,1	0,78	0,3	-0,48
Хол заг	0,68	0,63	0,68	0,05	0,55	0,42	-0,13
Тригліц	0,76	0,42	0,7	0,28	0,39	0,45	0,06
ЛПВЩ	0,68	0,06	0,09	0,03	0,05	0,08	0,03
ХЛНЩ	0,88	0,19	0,14	-0,05	0,14	0,14	0
ЛпДНЩ	0,88	0,23	0,28	0,05	0,19	0,25	0,06
Коеф атерог	0,8	0,27	0,36	0,09	0,28	0,16	-0,12
Кальцій іонізов	0,52	0,06	0,05	-0,01	0,05	0	-0,05
Лептин сиров	0,92	0,3	0,23	-0,07	0,27	0,2	-0,07

Значення, отримані в « Δ_2 » показують ймовірність того, що в результаті лікування M2 покращився показник, якщо значення позитивне, і відповідно погіршився, якщо значення негативне.

Значення, отримані в « Δ_3 » показують ймовірність того, що в результаті лікування M3 покращився показник, якщо значення позитивне, і відповідно погіршився, якщо значення негативне.

Наприклад, показник «САТ» покращився лікуванням M2 Різниця 0,51, а при M3 0,44; а показник «ДАТ» покращився лікуванням M3 0,39, а при M2 0,33.

Є певні показники, які погіршились після обох методик лікування. Наприклад: «Глюкоза», «Кальцій іоніз», «Лептин», а значить необхідні додаткові дослідження.

Висновки. У статті представлена інформаційна технологія опрацювання медичних показників на основі методу гістограмного аналізу. Гістограмний аналіз є важливим інструментом в медичній статистиці й дослідженнях. Цей метод дозволяє візуально і кількісно оцінити розподіл даних, що є корисним для виявлення відхилень, асиметрії та інших характеристик розподілів.

Аналіз показників, отриманих двома різними методиками лікування, показав як позитивні зміни, так і негативні, що потребують подальших досліджень. Особливістю запропонованої інформаційної технології є те, що вона дає змогу оцінити стан показника загалом, без прив'язки до конкретного пацієнта.

Список використаних джерел

1. Метод статистичного оцінювання результатів клінічних досліджень при лікуванні артеріальної гіпертензії та ожиріння / В. Єременко, О. Монченко, С. Корчева, Л. Чубко // Технічні науки та технології: науковий журнал. – 2023. – № 2 (32). – С. 164.
2. Мельникова, Н. І. Розроблення інформаційної технології опрацювання персоналізованих медичних даних / Н. І. Мельникова, О. Б. Вовк, Т. О. Дубінець // Вісник Національного університету "Львівська політехніка". Серія : Інформаційні системи та мережі. – 2015. – № 814. – С. 90-99.
3. Discretization from Data Streams: Applications to Histograms and Data Mining / M. Afonso-Dias, J. Simoes, C. Pinto // *Gis/spatial analysis in fishery and aquatic sciences : Proceedings 2th International Symposium on GIS/Spatial Analysis in Fishery and Aquatic Sciences.* – Saitama, 2014. – Pp. 323-340.
4. Піцун, О. Адаптивний метод попередньої обробки гістологічних та цитологічних зображень / О. Піцун // Вісник НУ «Львівська політехніка». Комп'ютерні науки та інформаційні технології. – 2017. – № 864. – С.111-118.
5. Mills, K. T. The global epidemiology of hypertension / K. T. Mills, A. Stefanescu, J. He // *NatRevNephrol.* – 2020. – № 16. – С. 223-237. DOI: <https://doi.org/10.1038/s41581-019-0244-2>.
6. WHO Global Health Observatory. Prevalence of raised blood pressure: situation and trends. [Electronic resource]. – Access mode: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en.
7. Малайчук, В. П. Математическая дефектоскопия : монография / В. П. Малайчук, А. В. Мозговой. – Днепропетровск : Системные технологии, 2005. – 180 с.

References

1. Eremenko, V., Monchenko, O., Korcheva, S., Chubko, L. (2023). Method of statistical evaluation of clinical trial results in the treatment of hypertension and obesity. *Technical sciences and technologies*, 2(32), 164.
2. Melnikova, N.I., Vovk, O., Dubinets, T. (2015). Development of information technology for processing personalized medical data. *Visnyk Natsionalnoho universytetu "Lvivska politekhnika". Seriya: Informatsiini systemy ta merezhi – Bulletin of the Lviv Polytechnic National University. Series: Information systems and networks*, 814, 90-99.
3. Afonso-Dias, M., Simoes, J., Pinto, C. (2014). Discretization from Data Streams: Applications to Histograms and Data Mining. *Gis/spatial analysis in fishery and aquatic sciences. Proceedings 2th International Symposium on GIS/Spatial Analysis in Fishery and Aquatic Sciences.* (pp. 323–340). Saitama, Japan.

4. Pitsun, O. (2017). Adaptive method of histological and cytological images pre-processing. *Visnyk NU «Lvivska politehnika». Kompiuterni nauky ta informatsiini tekhnologii – Bulletin of Lviv Polytechnic National University Computer Science and Information Technology*, 864, 111-118.
5. Mills, K.T., Stefanescu, A. & He, J. (2020). The global epidemiology of hypertension. *NatRevNephrol*, 16, 223–237. <https://doi.org/10.1038/s41581-019-0244-2>.
6. WHO Global Health Observatory. Prevalence of raised blood pressure: situation and trends. http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/.
7. Malaychuk, V.P., Mozgovoy, A.V. (2005). *Matematycheskaia defektoskopiia [Mathematical Defectoscopy]*. System Technologies.

Отримано 16.02.2024

UDC 616-085:311.218

Volodymyr Yeremenko¹, Olena Monchenko², Taras Monchenko³, Valentyna Kucherenko⁴

¹Doctor of Technical Science, Chief of Department of Information Measuring Technology
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Kyiv, Ukraine)
E-mail: nau_307@ukr.net. **ORCID:** <https://orcid.org/0000-0002-4330-7518>. **Scopus Author ID:** 56736964700

²PhD in Technical Sciences, Associate Professor Department of Biocybernetics and Aerospace Medicine
National Aviation University (Kyiv, Ukraine)
E-mail: monchenko_olena@ukr.net. **ORCID:** <https://orcid.org/0000-0002-8248-5704>
ResearcherID: AAN-1910-2020. **ID Scopus:** 56422110000

³Student, Department of Acoustic and Multimedia Electronic Systems
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Kyiv, Ukraine)
E-mail: mtm082220-ames27@ill.kpi.ua. **ORCID:** <https://orcid.org/0009-0002-3294-7023>

⁴PhD in Technical Sciences, Associate Professor Department of Biocybernetics and Aerospace Medicine
National Aviation University (Kyiv, Ukraine)
E-mail: valentyna.kucherenko@npp.nau.edu.ua. **ORCID:** <https://orcid.org/0000-0003-1009-8729>

INFORMATION TECHNOLOGY FOR PROCESSING MEDICAL INDICATORS

The article presents a new information technology for processing medical indicators, which is based on the method of histogram analysis. Histogram analysis, known as a key tool in medical statistics, allows visually and quantitatively evaluate the distribution of data, which becomes important for detecting asymmetry, deviations and other characteristics of the indicator base.

The analysis of the existing algorithms for presenting and processing medical information shows that they do not fully satisfy the requirements for solving problems that require complex logical conclusions, taking into account the incompleteness and inconsistencies of the input data. To solve this problem, it is necessary to look for new technologies for processing medical data.

It is proposed to use a histogram analysis of each indicator for the studied groups: a group of healthy patients (M1), a group of sick patients who received DASH treatment (M2), and a group of sick patients who were treated by the standard method (M3).

The purpose of the article is to develop an information technology for the analysis of the probability of changes in the values of medical indicators in the course of the chosen method of treatment.

The mathematical justification and algorithm of information technology for calculating medical indicators using the histogram method have been developed, fragments of normalized histograms and the obtained results are given. Information technology for processing medical indicators based on the method of histogram analysis is presented.

Histogram analysis is an important tool in medical statistics and research. This method allows visually and quantitatively evaluate the distribution of data, which is useful for detecting deviations, asymmetry and other characteristics of distributions.

Analysis of the indicators obtained by two different treatment methods showed both positive and negative changes, which require further research. A feature of the proposed information technology is that it allows assessing the state of the indicator as a whole, without reference to a specific patient.

Keywords: histogram analysis; information technology; medical indicators.

Fig.: 3. Table: 1. References: 7.