

О.В. Шкуліпа¹, Б.В. Рубльов¹, д.фіз.-мат. н.,
М.Г. Проданчук², д.мед.н., Б.С. Шейман³, д.мед.н.

ОЦІНКА ЕФЕКТИВНОСТІ ЗАСТОСУВАННЯ ЕЛІПТИЧНОЇ МАТЕМАТИЧНОЇ МОДЕЛІ ДЛЯ РОЗПІЗНАВАННЯ ЕТІОПАТОГЕНЕТИЧНОГО ЧИННИКА ЗАХВОРЮВАННЯ В ГРАФІЧНИХ ОБРАЗАХ ЕНДОТОКСЕМІЇ

¹Київський національний університет імені Тараса Шевченка,
²Інститут екології та токсикології ім. Л.І.Медведя МОЗ України,
³Національна дитяча спеціалізована лікарня "Охматдит" МОЗ України

Дослідження, розпочаті в 2000 році в різних країнах і присвячені вивченню процесів експресії генів під впливом різних чинників екзогенної й ендогенної природи, створили передумови для виникнення нового напрямку медицини, який отримав назву "токсикогеноміки". Значущість розвитку цього напрямку неможливо применшити, адже кожний медик обізнаний з термінами: нефрогеноміка, гепатогеноміка, токсикогеноміка та інші.

Прогностичним стало висловлювання академіка РАМН А.І. Арчакова — "Очевидно, содружество трех разделов — геномики, протеомики и биоинформатики — будет олицетворять будущее наук о жизни в XXI веке".

У 2003 році завершився міжнародний проект "Геном людини", а повна розшифровка геному людини відбулася в 2002 році, тобто через 50 років від винаходу подвійної спіралі молекул спадковості людини (ДНК). Тепер ми знаємо, що в людини не мільйон генів, а лише 30 тисяч, від них залежить наш розвиток від статевої клітини до дорослої форми. Ці роботи завершилися доказом того, що життєдіяльність організму людини супроводжується взаємопов'язаними процесами від експресії генів аж до метаболічних реакцій, які відбуваються на рівні органів, систем, клітин.

Поступово стало зрозумілим: всі організми мають гени, що продукують специфічні білки, у яких основною функцією є попередження або виправлення хімічних або фізичних ушкоджень ДНК. Досягнення науки щодо участі специфічних білків у реалізації клітинної відповіді на дію пошкоджуючих факторів довкілля сформувало новий напрямок у біології та медицині, який отримав назву "токсикопротеоміка". Його появу в сучасних медико-біологічних дослідженнях можна вважати одним із видатних досягнень у галузі геномної біотехнології ХХ сторіччя. На сучасному етапі вже відомі та певною мірою зрозумілі процеси нормальної генної послідовності, структура трансльованих білків та рівні активності.

Результати протеомних досліджень надають надзвичайно важливу інформацію про фізіологічні функції на субклітинному, клітинному, органному й системному рівнях, що відкриває нові горизонти для розуміння молекулярних механізмів дії ліків та розвитку захворювань, ефектів дії на організм факторів довкілля.

Досягнення науки в галузі токсикогеноміки, імуногенетики й протеоміки дозволяють розглядати алгоритм формування захворюван-

ня й реалізації токсикозу як послідовну зміну взаємопов'язаних процесів, від етапу проникнення етіологічного чинника в організм аж до виникнення інтоксикаційного синдрому — клінічної маніфестації ендотоксикозу. Відомо, що успіх у лікуванні різних захворювань у більшості залежить від усунення, в першу чергу, причини, що викликала це захворювання, тобто — від ефективності етіотропної терапії. Якщо взяти до уваги, що етіологічний чинник та етіопатогенетичні механізми формування хвороби, з одного боку, та накопичення в кров'яному руслі автоагресивних речовин ендо- або екзогенного походження (токсемія), з іншого, — ланки єдиного патологічного процесу, то стає зрозумілою роль етіотропної терапії в лікуванні токсикозу.

Цілком природно, оперативна верифікація етіологічного чинника, що викликав розвиток токсемії, дозволила б у максимально короткі терміни знизити або ліквідувати його активність, призупинити синтез ендотоксинів, або усунути ускладнення, які виникають під час перебігу захворювання та пов'язані з розвитком токсикозу. Проте слід урахувати, що низка захворювань може бути викликана поліетіологічними причинами, а спроба їхньої ідентифікації є вкрай складним та неперспективним завданням для прийняття оперативних клінічних рішень. Крім цього, в процесі розвитку основного захворювання може відбуватися активація хронічних вогнищ вірусно-паразитарної та бактеріальної інфекції, виникати гостра або поглиблюватися хронічна недостатність фізіологічних систем детоксикації. Ці порушення неминуче призведуть до накопичення в крові автоагресивних речовин (АР) з різними токсикометричними параметрами й характеристиками. Саме тому дослідження пошкоджуючих властивостей АР, які циркулюють у кров'яному руслі, їхньої етіологічної приналежності, є важливим для лікування токсикозу в цілому, а також для індивідуалізації показань щодо призначення етіотропної та детоксикаційної терапії, зокрема.

Одним із сучасних підходів до діагностики етіологічних чинників є розпізнавання останніх у "графічному образі" токсемії. Як приклад можна навести посилання на таку

асоціацію. Розглядаючи відому картину Леонардо да Вінчі "Джоконда", ні в кого не виникає сумнівів, що зображено людину, ця людина — жіночої статі, представник білої раси. А більш глибокі знання дозволяють сказати більше: це портрет Мони Лізи, назвати прізвище художника, у якому році картину написано, а також те, що вона знаходиться у Франції, у Луврі та ін. Уявляється вочевидь, що візуальна оцінка графічного образу є достатньо ефективним, інформативним та оперативним засобом ідентифікації об'єкту. Повною мірою це стосується й клінічної токсикології, де ми зробили спробу відтворення характерних графічних зображень токсемії. Для цього було здійснено комплексні дослідження токсикометричних параметрів останньої, а отримані результати розділили на групи (бібліотеки) залежно від ідентифікованої традиційними загальноновизнаними засобами етіопатогенетичної причини захворювання. Таким чином, була створена бібліотека графічних образів-еталонів токсемії, що виникає за різних етіологічних чинників.

Ідея розробки методики розпізнавання етіологічного чинника за "графічним образом" токсемії базується на попередньому визначенні токсикометричних параметрів токсемії у хворого та створення її графічного образу. Потім проводять порівняння параметрів графічного образу токсемії у хворого з такими в різних бібліотеках. Ідентифікація етіологічного чинника токсемії у хворого відбувається на підставі обрахунку результатів математичного аналізу за максимальною кількістю достовірних збігів між параметрами графічного образу токсемії, що досліджується, та відповідним графічним образом-еталоном із бібліотеки з відомим етіологічним чинником.

Є цілком зрозумілим, що процес ідентифікації етіологічного чинника, його інформативність, оперативність та ефективність у повній мірі залежить від обраної для аналізу математичної моделі. Саме пошуку найбільш удосконаленої та ефективної математичної моделі для розпізнавання етіопатогенетичних чинників у графічному образі токсемії й присвячена ця наукова робота.

Метою наукової роботи є розробка алгоритму кластеризації да-

них шляхом побудови математичної моделі зі знаходженням опуклої оболонки серед множини точок на площині, побудови еліпсів мінімальної площі, що апроксимують шукані кластери, для визначення етіопатогенетичного чинника токсемії в пацієнтів із різними захворюваннями.

Матеріали та методи дослідження. Для побудови математичної моделі були використані токсикометричні характеристики ендотоксемії, яка супроводжувала перебіг різних захворювань у 500 пацієнтів. З них за допомогою традиційних загальноновизнаних методів дослідження було встановлено етіопатогенетичні чинники ендотоксемії: аномалії печінки — 34; аномалії нирок — 30; аутоімунні/аутоалергічні реакції — 158; бактеріальні збудники — 114; вірусні/паразитарні збудники — 85; деструкція тканин різного походження — 19; локальні гіперпластичні процеси в тканинах (новоутворення) — 14; системні запальні реакції (SIRS) — 23; отрути екзогенного походження — 23. Систематизовані за етіопатогенетичними чинниками токсикометричні дані ендотоксемії склали основні класи бібліотеки (еталони графічних образів токсемії). 45 графічних образів ендотоксемії (по 5 графічних образів у класі за кожним етіопатогенетичним чинником) становили контрольну групу, яку використовували для перевірки отриманих результатів дослідження.

З математичної точки зору задача визначення етіологічного фактора на підставі особливостей прояву ендотоксемії зводиться до наступної математичної моделі:

Для заданого графічного образу ендотоксемії $x = (x_1, \dots, x_n)$ необхідно

визначити етіологічний чинник виникнення хвороби.

Уведемо позначення, що будуть використовуватись в наступному аналізі:

- g - кількість класів, тобто кількість етіологій, $g=2$.
- n_k - кількість спостережень в k -тій етіології.
- n - загальна кількість спостережень, $n=65$.
- x_{ikm} - величина змінної i для m -го спостереження в k -тій етіології $i=1, \dots, 29$.

Опис системи. В основу алгоритму визначення етіологічного чинника ендотоксемії покладено метод кластеризації даних, шляхом знаходження опуклої оболонки серед множини точок на площині та подальшої побудови еліпсів мінімальної площі (ЕМП), що апроксимують шукані кластери. Відомо, якщо дані розміщені за нормальним розподілом, то їхня множина точок буде апроксимуватися еліпсом. З метою усунення впливу помилкових даних для побудови кластерів нами були відібрані лише ті токсикометричні показники, які найбільше впливають на кластеризацію етіологій. Для цього на початку був проведений аналіз опуклих оболонок для кожної пари етіологічних чинників, а потім застосовано побудову еліпсу мінімальної площі для однієї з пари етіологій.

Алгоритм побудови опуклої оболонки. Для побудови опуклої оболонки застосовували алгоритм Грехема [1]. Суть алгоритму полягає в однократному перегляді впорядкованої послідовності точок, у процесі якого вилучаються внутрішні точки. Точки, що лишилися та які подано у відповідному порядку, є вершинами оболонки (рис. 1).

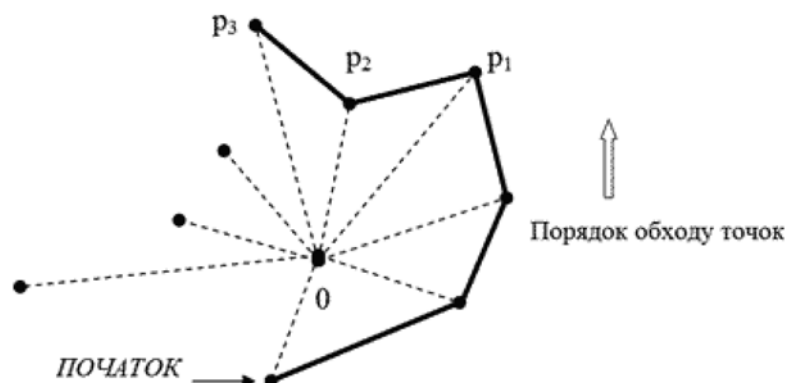


Рис. 1. Початок обходу точок за методом Грехема (вершина p_2 вилучається, якщо кут $p_1p_2p_3$ виявляється угнутим)

Перегляд починається з точки, яку позначено як "ПОЧАТОК". В якості цієї точки можна брати "найправішу" із найменшою ординатою з даної множини, яка напевне, є вершиною опуклої оболонки. Трійки послідовних точок багатократно перевіряють у порядку обходу проти годинникової стрілки з метою визначення, чи утворюють вони кут, більший або рівний p . Якщо внутрішній кут $p_1p_2p_3 \geq p$, $p_1p_2p_3$ утворюють "правий поворот" ($\Delta > 0$), інакше $p_1p_2p_3$ утворюють "лівий поворот" ($\Delta < 0$). Δ визначається за наступною формулою:

$$\Delta = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

Залежно від результатів перевірки кута, який утворюється трійкою вершин, можливі два варіанти продовження перегляду:

- $p_1p_2p_3$ утворюють "правий поворот" — вилучити вершину p_2 і перевірити трійку $p_1p_3p_4$;
- $p_1p_2p_3$ утворюють "лівий поворот" — продовжити перегляд, перейшовши до трійки $p_2p_3p_4$.

Алгоритм закінчує роботу, коли прохід точок повертається на початок списку.

Алгоритм побудови еліпсу мінімальної площі (ЕМП) для довільного багатокутника. Для побудови ЕМП ми застосовували наступний алгоритм [2]:

1. Розглядаємо всі трикутники вигляду $A_iA_jA_k$, де $1 \leq i < j < k \leq n$ (n — загальна кількість точок, утворених вершинами багатокутника M). Серед них знаходимо трикутник $A_iA_jA_k$ найбільшої площі. Для цього трикутника будуємо еліпс мінімальної площі e_{ijk} . Якщо виконується включення "Помилка! Закладку не визначено", в такому разі шуканий ЕМП для всього багатокутника M і алгоритм побудови ЕМП завершується; якщо $M \not\subseteq e_{ijk}$ — то переходимо до виконання наступного кроку.
2. Розглядаємо всі чотирикутники вигляду $A_iA_jA_kA_l$, де $1 \leq i < j < k < l \leq n$, де (n — загальна кількість точок, що утворюються вершинами багатокутника M). Для кожного з них перевіряємо умову $M \subseteq e_{ijkl}$. Де e_{ijkl} є ЕМП для відповідного чотирикутника (i серед усіх ЕМП

розглядаємо лише ті, що містять усі чотири його вершини). Якщо ця умова виконується для одного із чотирикутників, то відповідний ЕМП буде ЕМП для всього багатокутника M , і алгоритм завершує роботу. Якщо включення $M \subseteq e_{ijkl}$ не виконується для жодного з еліпсів, то переходимо до виконання останнього кроку алгоритму.

3. Розглядаємо всі можливі п'ятикутники $A_iA_jA_kA_lA_m$, де $1 \leq i < j < k < l < m \leq n$, що утворюються вершинами багатокутника M , і для кожного з них будуємо відповідний ЕМП e_{ijklm} , що містить усі вершини відповідного п'ятикутника, а тому він існує єдиний, що містить багатокутник M . Цей ЕМП і є шуканим, і алгоритм завершує роботу.

Алгоритм розпізнавання етіопатогенетичного чинника токсемії за допомогою еліптичного підходу на прикладі факторів "Аномалії печінки" та "Аномалії нирок". Для визначення тих токсикометричних параметрів, які є значимими при розділенні етіологічних чинників токсемії (на прикладі "Аномалії печінки" та "Аномалії нирок") проведено попарне їхнє групування в двовимірному просторі.

У табл. 1 зображена матриця з токсикометричними параметрами, які групуються.

Кількість таких комбінацій буде дорівнювати

$$C_n^2 = \frac{n!}{2!(n-2)!} = C_{29}^2 = 406$$

Використання попарного аналізу є зручним, адже він може бути проведений у двовимірному просторі. За допомогою методу Грехема, у кожному з випадків $i_{l,m}$ може бути побудована опукла оболонка для кожного із двох етіопатогенетичних чинників ендотоксемії.

Приклади побудови опуклих оболонок наведено на рис. 2, 3.

На рис. 2 зображено результат кластеризації етіопатогенетичних чинників ендотоксемії за токсикометричними показниками НСТ-Гсп. та НСТ-Гінд. Отримані дані свідчать про досить проблематичну кластеризацію, адже, по-перше, опуклі оболонки перетинаються; по-друге, хворий, у якого ендотоксемія спричинена етіопатогенетичним чинником "Аномалії печінки", потрапляє в опуклу оболонку, що створена етіопатогенетичним чинником "Аномалії нирок".

На рис. 3 зображена протилежна

Таблиця 1

Матриця порівнянь

Токсикометричні параметри	НСТ-Гсп.	НСТ-Гінд.	НСТ-М сп.	.	.	.	ЦАЛс-МВ	ЦІК	КрЛг
НСТ-Гсп.	0			.	.	.			
НСТ-Гінд.	0	0		.	.	.			
НСТ-М сп.	0	0	0	.	.	.			
.
.
.
ЦАЛс-МВ	0	0	0	.	.	.	0		
ЦІК	0	0	0	.	.	.	0	0	
КрЛг	0	0	0	.	.	.	0	0	0

Примітка:

1. НСТ-Гсп. - спонтанний НСТ-тест гранулоцитів.
2. НСТ-Гінд. - індукований НСТ-тест гранулоцитів.
3. НСТ-Мсп. - спонтанний НСТ-тест моноцитів.
4. ЦАЛс-МВ - цитолітична активність вільноциркулюючих ендотоксинів із розміром молекул < 10 нм.
5. ЦІК - циркулюючі імунні комплекси.
6. КрЛг - криолабільні глобуліни.
7. $i_{1,2}$ означає, що етіопатогенетичні чинники "Аномалії печінки" та "Аномалії нирок" порівнюються за токсикометричними параметрами "НСТ-Гсп." та "НСТ-Гінд."
8. "0" означає, що такі групи відкидаються з аналізу, через їхню симетричність.

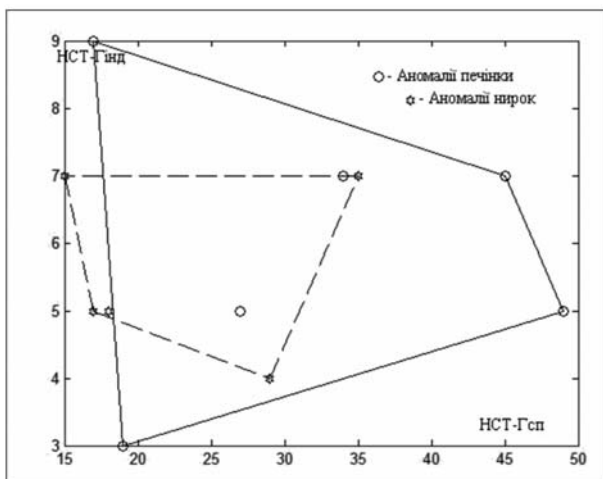


Рис. 2 Опуклі оболонки у випадку НСТ-Гсп. та НСТ-Гінд.
Примітка:
1. НСТ-Гсп. – спонтанний НСТ-тест гранулоцитів.
2. НСТ-Гінд. – індукований НСТ-тест гранулоцитів.

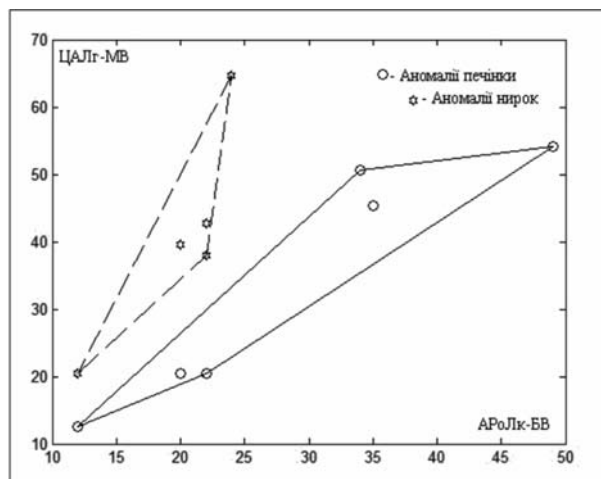


Рис. 3 Опуклі оболонки у випадку АРОЛк-БВ та ЦАЛг-МВ.
Примітка:
1. АРОЛк-БВ – автоімунна активність цитомембранних ендотоксинів із розміром часток > 200 нм.
2. ЦАЛг-МВ – цитолітична активність глобулін-асоційованих ендотоксинів із розміром молекул < 10 нм.

ситуація, коли за допомогою токсикометричних параметрів АРОЛк-БВ та ЦАЛг-МВ можна провести класифікацію.

Після аналізу отриманих результатів побудови опуклої оболонки, залишились тільки ті токсикометричні показники, в яких опуклі оболонки або не перетинались, або їх можна розділити за допомогою кривої другого порядку $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{21}y + a_{33} = 0$ (рис. 4, 5).

Таким чином, нами було отримано 33 пари. Провідним показником, який розділяє етіопатогенетичні чинники (в нашому прикладі), був "АРОЛ-г" (його рівень був визначальним для розділення-розпізнавання етіопатогенетичного чинника в 51,5 % випадків); наступним за ієрархією був показник "ЦАЛс-МВ"

(його рівень був визначальним у 15,15 % випадків).

Для подальшої кластеризації будували ЕМП для одного з етіопатогенетичних чинників та проводили перевірку отриманих результатів на контрольній вибірці хворих (рис. 6, 7).

У результаті проведених досліджень нами були отримані рівняння еліпсів, що кластеризують два токсикометричні показники. Для токсикометричних показників НСТ-Гсп. та АРОЛ-г рівняння еліпсу мінімальної площі набуває вигляду: $f(x,y) = -0.0002155x^2 - 5.7291e - 005xy - 0.00017068y^2 + 0.017827x + 0.02379y - 1 = 0$

З метою перевірки отриманого нами рівняння еліпсу мінімальної площі для показників НСТ-Гсп. та АРОЛ-г були проведені тестові дослідження з використанням токсикометричних параметрів ендотоксемії в пацієнтів контрольної групи (табл. 2).

Результати тестових досліджень розпізнавання етіопатогенетичного чинника ендотоксемії за параметрами НСТ-Гсп. та АРОЛ-г

Таблиця 2

Результати тестових досліджень розпізнавання етіопатогенетичного чинника ендотоксемії за параметрами НСТ-Гсп. та АРОЛ-г

Етіопатогенетичний чинник ендотоксемії	Кількість випадків правильного розпізнавання (%)
Аномалії печінки	91,17
Аномалії нирок	60,00

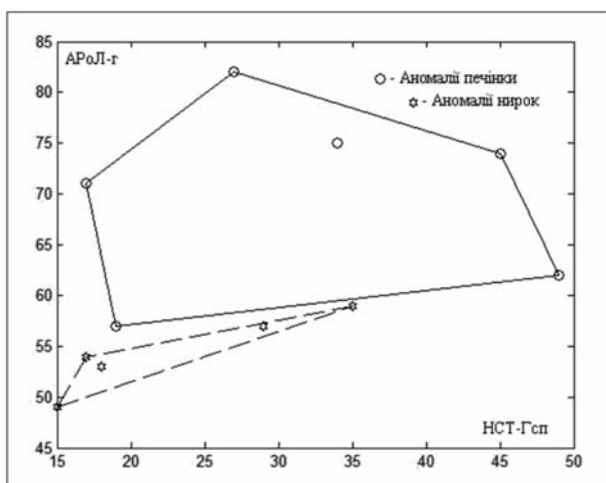


Рис. 4 Токсикометричні показники (АРОЛ-г та НСТ-Гсп.), що розділяють етіопатогенетичні чинники

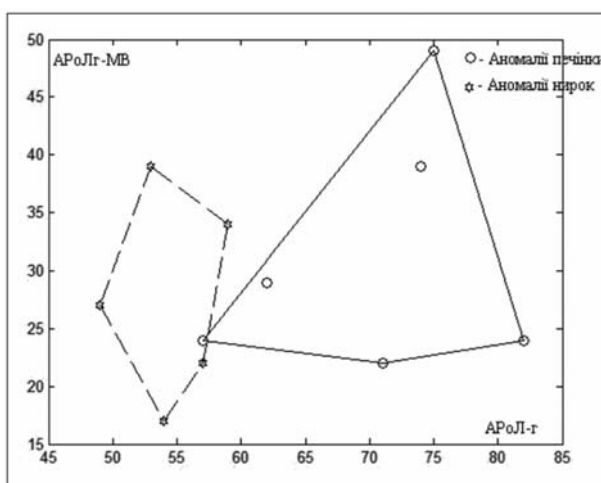


Рис. 5 Токсикометричні показники, що розділяють етіопатогенетичні чинники

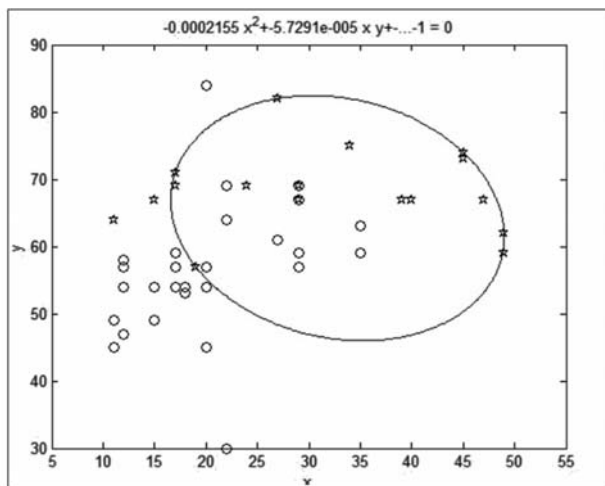


Рис. 6 Кластеризація даних за токсикометричними показниками НСТ-Гсп та АРоЛ-г за допомогою еліптичного підходу.

Примітка (тут та надалі):

- 1) * – параметри ендотоксемії у хворих з етіопатогенетичним чинником "Аномалії печінки";
- 2) ° – параметри ендотоксемії у хворих з етіопатогенетичним чинником "Аномалії нирок".

Для випадку токсикометричних показників НСТ-Гінд та АРоЛ-г рівняння ЕМП набуває вигляду: $f(x,y) = -9.2824e - 005x^2 - 0.00012723xy - 7.3637e - 005y^2 + 0.018738x + 0.015551y - 1 = 0$

Результати перевірки алгоритму на тестових даних для НСТ-Гінд та АРоЛ-г надано в табл. 3:

Результат роботи алгоритму — знайдене рівняння еліпса, що класифікує етіології: $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{21}y + a_{33} = 0$.

Тобто, якщо токсикометричні параметри хворого потрапляють усередину еліпса, то ендотоксемія за етіопатогенетичним чинником відноситься до "Аномалії печінки";

якщо зовні еліпса — до "Аномалії нирок". Важливим є встановлений результат, який вказує на існування декількох пар токсикометричних показників, за якими можна з імовірністю понад 70 % визначити етіопатогенетичний чинник виникнення ендотоксикозу. Можна припустити, що пошук оптимальної математичної моделі розпізнавання етіопатогенетичного чинника ендотоксемії повинно відкрити нові підходи до діагностики в медицині та сприяти оперативному застосуванню методів етіотропної терапії в лікуванні різних захворювань.

Висновки.

1. Для розпізнавання етіопатогене-

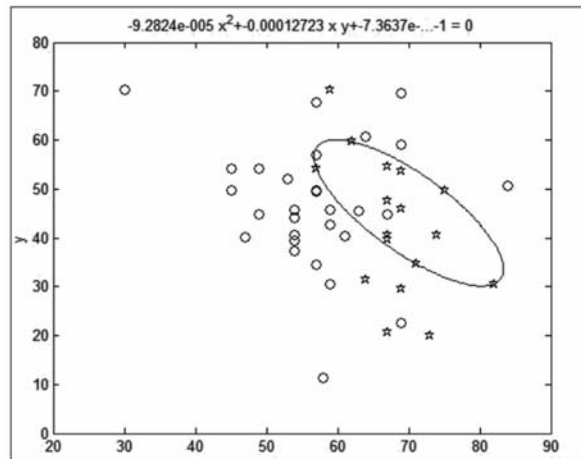


Рис. 7 Кластеризація за токсикометричними показниками АРоЛ-г та ЦАЛа-МВ за допомогою еліптичного підходу

Таблиця 3

Результати тестових досліджень розпізнавання етіопатогенетичного чинника ендотоксемії за параметрами НСТ-Гінд та АРоЛ-г

Етіопатогенетичний чинник ендотоксемії	Кількість випадків правильного розпізнавання
Аномалії печінки	97,05
Аномалії нирок	66,70

тичного чинника захворювання в графічних образах ендотоксемії можна застосовувати еліптичну математичну модель.

2. Рівняння еліпсів мінімальної площі для різних пар токсикометричних параметрів дозволяє з високою точністю (від 60,00 % до 97,05 % випадків) визначити етіопатогенетичний чинник ендотоксемії у хворого.

Роботи з пошуку оптимальної математичної моделі розпізнавання етіопатогенетичного чинника ендотоксемії в пацієнтів із різними захворюваннями триває.

ЛІТЕРАТУРА

1. Препарата Ф. Вычислительная геометрия: Введение. / Ф.Препарата, М. Шеймос, Мир, 1989. — 478 с..
2. Рубльов Б.В. Геометричні властивості еліпса мінімальної площі та деякі суміжні питання / Б.В. Рубльов, Ю.І. Петунін, Ю.Ю. Милейко / К.:Київський університет, 2000. — 73 с.

Е.В.Шкуліпа, Б.В.Рубльов, Н.Г.Проданчук, Б.С.Шейман

ОЦЕНКА ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ ЭЛЛИПТИЧЕСКОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ДЛЯ РАСПОЗНАВАНИЯ ЭТИОПАТОГЕНЕТИЧЕСКОГО ФАКТОРА ЗАБОЛЕВАНИЯ В ГРАФИЧЕСКИХ ОБРАЗАХ ЭНДОТОКСЕМИИ

В работе продемонстрирована возможность применения эллиптической математической модели для распознавания этиопатогенетических факторов эндотоксемии у больных. Рассмотрена эффективность использования предложенной модели для диагностики заболевания.

Ключевые слова: болезни, эндотоксемия, этиопатогенез, диагностика, математическая модель, кластерный анализ, эллипс минимальной площади

O.V. Shkulipa, B.V. Rublyov, N.G. Prodanchuk, B.S. Sheyman

ESTIMATION OF EFFICIENCY OF APPLICATION OF ELLIPTIC MATHEMATICAL MODEL FOR RECOGNITION OF ETIOPATHOGENETICS FACTORS OF ENDOTOXEMIA

The possibility of application of elliptic mathematical model is in-process shown for recognition of etiopathogenetics factors of endotoxemia in patients. Efficiency of the use of the offered model is considered for diagnostics of disease.

Key words: illnesses, endotoxemia, etiopathogenesis, diagnostics, mathematical model, cluster analysis, ellipse of minimum area

Надійшла до редакції 14.01.10