

**МАТЕМАТИКО-СТАТИСТИЧНИЙ АНАЛІЗ ПРОГНОЗУВАННЯ СТУПЕНЯ ТЯЖКОСТІ СТАНУ ХВОРИХ НА ХРОНІЧНЕ ОБСТРУКТИВНЕ ЗАХВОРЮВАННЯ ЛЕГЕНЬ З СУПУТНЬОЮ ГІПЕРТОНІЧНОЮ ХВОРОБОЮ**

Харківський національний медичний університет (м. Харків)

istominaolya@yandex.ru

Робота є фрагментом НДР «Удосконалення оцінки кардіоваскулярного ризику за хронічного обструктивного захворювання легень», № державної реєстрації 0116U004989.

**Вступ.** Хронічне обструктивне захворювання легень (ХОЗЛ) в поєднанні з гіпертонічною хворобою (ГХ) представляє одне з частих коморбідних станів в клініці внутрішніх хвороб. У хворих на ХОЗЛ враховуючи запальний генез і його системні прояви рано порушується функція ендотелію, як й при ГХ за рахунок гемодинамічного навантаження [1,5]. До теперішнього часу не проведено комплексне дослідження функціонального стану ендотелію у хворих на ХОЗЛ в поєднанні з ГХ та відсутній чіткий алгоритм прогнозування стану тяжкості у таких хворих.

**Метою роботи** була розробка алгоритму індивідуального прогнозування ступеня тяжкості стану у хворих на ХОЗЛ із супутньою ГХ на основі даних біохімічного дослідження сироватки крові.

**Об'єкт і методи дослідження.** Відповідно до програми дослідження було обстежено 100 пацієнтів, які знаходилися на обстеженні та лікуванні в клініці НДІ гігієни праці та професійних захворювань ХНМУ – клінічній базі кафедри внутрішніх та професійних хвороб. Всі пацієнти були розподілені на групи: до першої увійшли хворі на ХОЗЛ з супутньою гіпертонічною хворобою (55 осіб – 35 чоловіків, 20 жінок) – основна група; до другої групи увійшли пацієнти з ізольованою ХОЗЛ (45 осіб – 34 чоловіка, 11 жінок) – група порівняння. Групу контролю склали 45 умовно здорових пацієнтів (25 чоловіків, 20 жінок). Середній вік в середньому складав  $55,4 \pm 10,7$  років і був співставним в усіх трьох групах. У всіх обстежених проводились такі дослідження: клінічний та біохімічний аналізи крові, визначались рівні катехоламінів, метаболітів та синтаз оксиду азоту, S-нітрозотіол, VEGF, рівні біогенних елементів, вазоактивних гормонів, пептидів та нейромедіаторів. Статистична обробка результатів дослідження проведена з використанням пакету програм «Statistica 6.0». Використано методи параметричної та непараметричної статистики, кластерного і дискримінантного аналізів [2,3,4,6,7].

**Результати дослідження та їх обговорення.** Згідно міжнародним рекомендаціям GOLD (2016) стадії тяжкості ХОЗЛ варіюють від I (A) — легка до IV (D) — вкрай важка. Однак, якщо у хворого наявність ХОЗЛ супроводжується іншими супутніми захворюваннями, серед яких на першому місці ГХ, то прогнозування стану хворого є досить складним завданням. Проблема полягає в тому, що перебіг за-

хворювання у хворих на ХОЗЛ навіть групи С, але без ГХ, може проходити менш важко, ніж перебіг захворювання В групи ХОЗЛ при наявності ГХ. Тому в загальному випадку ми маємо (4x2) варіантів перебігу основного захворювання. Всі методи класифікації таких випадків дуже трудомісткі [3] і засновані на оцінці відстаней між об'єктами (в нашому випадку — пацієнти) в n-вимірному просторі використовуваних змінних (в нашому випадку — біохімічні показники).

Тому в якості першого кроку ми використовували методи кластерного аналізу для знаходження однорідних груп. Попередньо між усіма вхідними показниками, значення яких стандартизувалися ( $n = 67$ ), методами кореляційно-регресивного аналізу були встановлені кореляційні залежності. Якщо значення коефіцієнта кореляції Спірмена між змінними перевищувало 0,7, то одна з змінних виключалася. В результаті були виділені 25 стандартизованих показників.

Було використано кластерний ієрархічний (деревовидний) алгоритм об'єднання об'єктів. Функцією відстані служила Евклідова метрика (Euclidean distances). Реалізовувався незважений центроїдний метод об'єднання в кластери, при якому відстань між двома кластерами визначається, як відстань між їх центрами тяжіння — деревоподібна кластеризація (tree dustering), що дозволило виявити з 100 пацієнтів дві однорідні групи, різні з яких-небудь причин (рис. 1).

Виділення саме 2-х кластерів було підтверджено методами дивізійного аналізу середніх показників кластерів (к-середніх). Вибір змінних (показників)

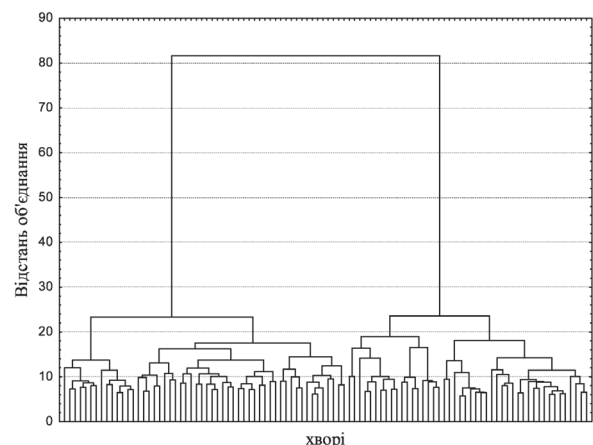
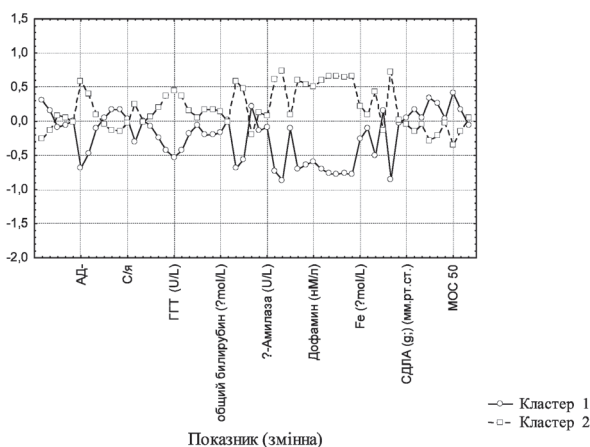


Рис. 1. Дендрограма розподілу спостережень за кластерами.

проводився на основі найкращого поділу кластерів за середнім значенням показників (рис. 2). Використання методу дивізійного аналізу k-середніх дозволило також отримати якісний склад кожного кластера, тобто віднести кожного хворого до відповідного кластера.



**Рис. 2. Стандартизовані середні показники двох кластерів.**

Також проведені на підставі тесту Манна-Уїтні парні порівняння даних двох кластерів між собою виявили істотні відмінності між ними практично за всіма біохімічними і фізіологічними показниками.

Отримані результати дозволяють зробити висновок про те, що кластери являють собою групи пацієнтів, що відрізняються між собою за ступенем тяжкості перебігу захворювання ХОЗЛ з урахуванням наявності ГХ, тобто поділ на кластери відповідає поділу на групи пацієнтів за ступенем тяжкості клінічного перебігу ХОЗЛ у поєднанні з ГБ: 1-й кластер — більш важкий ступінь стану хворого; 2-й кластер — середній ступінь тяжкості стану хворого.

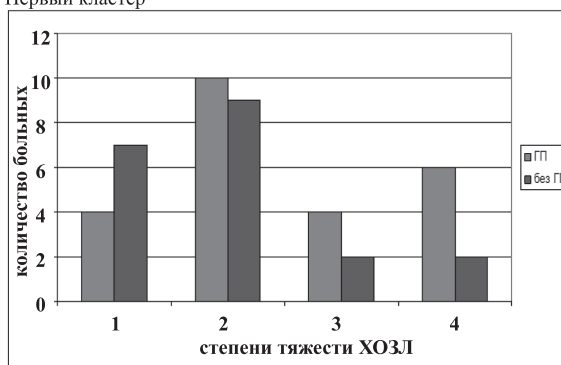
При цьому інформативними і прогностично значущими показниками для хворих з поєднаним перебігом ХОЗЛ і ГХ є маркери ендотеліальної дисфункції — вміст нітратів, нітритів, S-нітрозотіолів, активності eNOS, iNOS; фактор ангіогенезу — VEGF; високі рівні в крові заліза, кальцію.

На **рисунку 3** представлені діаграми розподілу по кластерам кількості пацієнтів з різними стадіями ХОЗЛ в поєднанні з ГБ і без неї. Як бачимо, 1-й кластер включає в себе більшу кількість хворих з важкими стадіями ХОЗЛ та супутньою ГБ (n = 14) в порівнянні з 2-м кластером (n = 9).

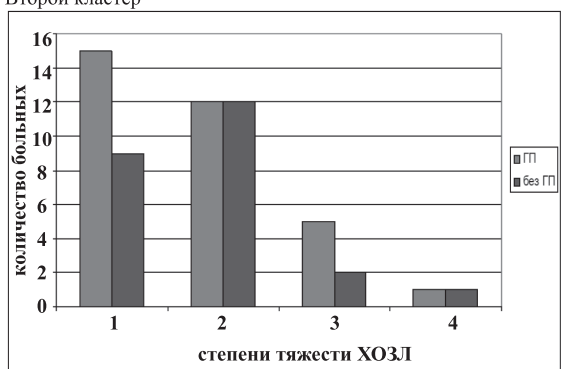
На підставі отриманої кластерної структури для прогнозування ступеня тяжкості стану хворих з ХОЗЛ і ГХ беручи до уваги числовий характер показників і їх закон розподілу було використано дискримінантний аналіз [4].

Відбір дискримінантних змінних з 25 вхідних стандартизованих показників, отриманих на етапі кластерного аналізу, проходив покроково методом виключення змінних відповідно до їх рівнем толерантності (міра надмірності змінних для класифікації).

Первый кластер



Второй кластер



**Рис. 3. Діаграма розподілу по кластерам хворих на ХОЗЛ у поєднанні з ГХ і без неї.**

Класифікація тяжкості стану будь-якого хворого з ХОЗЛ проводилася на підставі класифікаційної функції для цього хворого. Хворий ставився до того кластера (групи тяжкості стану), для якого чисельне значення класифікаційної функції максимально. Класифікаційні функції обчислювалися за формулами:

$$Y_{1\text{кластер}} = -113,32 + 1,223 \cdot \text{«Hb»} + 0,043 \cdot \text{«VEGF»} + 3,18 \cdot \text{«iNOS»} + 0,612 \cdot \text{«NO}_2\text{»} + 2,38 \cdot \text{«S-нітрозотіол»} + 0,043 \cdot \text{«NO}_3\text{»};$$

$$Y_{2\text{кластер}} = -123,26 + 1,27 \cdot \text{«Hb»} + 0,032 \cdot \text{«AcT»} + 0,05 \cdot \text{«VEGF»} + 3,6 \cdot \text{«iNOS»} + 0,537 \cdot \text{«NO}_2\text{»} + 4 \cdot \text{«S-нітрозотіол»} + 0,008 \cdot \text{«NO}_3\text{»}.$$

Про точність дискримінантної моделі та її ефективність можна судити за результатами апостеріорної класифікації. Тобто передбачається, що ступінь тяжкості захворювання нам не відома, і ми намагаємося її спрогнозувати на основі отриманих класифікаційних функцій. Результати такого аналізу в цілому по всій вибірці і окремо по групах дають дуже високу точність класифікації: по всій вибірці (n = 98) — 83,3%; по 1-й групі (1-й кластер, n = 44) — 81,8%; по 2-й групі (2-й кластер, n = 54) — 85,9% (**табл.**).

**Висновки.** В результаті проведеного математико-статистичного аналізу прогнозування ступеня тяжкості стану хворих з поєднаним перебігом ХОЗЛ і ГХ встановлено наступне:

1. Використання кластерного (деревовидного) алгоритму об'єднання об'єктів дозволило виявити з 100 пацієнтів 2 однорідні групи — 2 кластера.

Таблиця.

**Матриця класифікації по 2 кластерам**

Група	Матриця класифікації		
	Відсоток правильних, %	Кластер 1	Кластер 2
Кластер 1	81,8	36	8
Кластер 2	85,9	8	46
Всього	83,7	44	54

2. Розподіл хворих за двома кластерами відповідає виділенню двох груп пацієнтів з різним ступенем тяжкості їх стану при поєднаному перебігу ХОЗЛ і ГХ: 1-й кластер — це група хворих з більш тяжким ступенем стану; 2-й кластер — група хворих із середнім ступенем тяжкості стану.

3. В оцінці ступеня тяжкості стану хворого з ХОЗЛ і ГХ прогностично значущими показниками є марке-

ри ендотеліальної дисфункції — вміст нітратів, нітритів, S-нітрозотіолів, активності eNOS, iNOS; фактор ангіогенезу — VEGF; високі рівні в крові заліза, кальцію.

4. Методами дискримінантного аналізу розроблено алгоритм індивідуального прогнозу ступеня тяжкості стану хворого з поєднаним перебігом ХОЗЛ і ГХ. Класифікація тяжкості стану будь-якого хворого з ХОЗЛ проводиться на підставі встановлених розрахункових класифікаційних функцій Y1 кластер і Y2 кластер для даного хворого. Хворий відноситься до того кластеру (групі тяжкості стану), для якого чисельне значення класифікаційної функції максимальне. Матриця класифікації за двома кластерами підтверджує її високу точність.

**Перспективи подальших досліджень.** У зв'язку з виявленням прогностично значущих показників ендотеліальної дисфункції у хворих на хронічне обструктивне захворювання легень в поєднанні з гіпертонічною хворобою набуває перспективи розробка нових фармакологічних підходів до їх інгібіції у даної когорти хворих.

**Література**

1. Аксенова Т.А. Эластические свойства сосудов у больных хронической обструктивной болезнью легких, коморбидной с артериальной гипертензией / Т.А. Аксенова, В.В. Горбунов, С.Ю. Царенок // Дальневосточный медицинский журнал. — 2013. — № 2. — С. 10-12.
2. Варакин А.Н. Статистические модели регрессионного типа в экологии и медицине / Под. ред. В.Н. Чуканова. — Екатеринбург, 2006. — С. 158.
3. Жамбю М. Иерархический кластер-анализ и соответствия / М. Жамбю. — М.: Финансы и статистика, 1988. — С. 342.
4. Ким Дж.-О. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мюллер, У.Р. Клекка. — Москва: Финансы и статистика, 1989. — 215 с.
5. Кузьминова Н.В. Функциональное состояние сосудистого эндотелия у больных гипертонической болезнью / Н.В. Кузьминова, В.К. Серкова // Український терапевтичний журнал. — 2008. — № 2. — С. 21-27.
6. Мандель И.Д. Кластерный анализ / И.Д. Мандель. — М.: Финансы и статистика, 1988. — 176 с.
7. Халафян А.А. Статистический анализ данных. 3-е изд.-е. / А.А. Халафян. — Москва: Бином-Пресс, 2007. — 512 с.

УДК [616.24-007.272-036.12+616.12-008.331.1]-018.74-008.6-07:57.083.3

**МАТЕМАТИКО-СТАТИСТИЧНИЙ АНАЛІЗ ПРОГНОЗУВАННЯ СТУПЕНЯ ТЯЖКОСТІ СТАНУ ХВОРИХ НА ХРОНІЧНЕ ОБСТРУКТИВНЕ ЗАХВОРЮВАННЯ ЛЕГЕНЬ З СУПУТНЬОЮ ГІПЕРТОНІЧНОЮ ХВОРОБОЮ**  
**Капустник В. А., Истомина О. В., Зайцева О. В.**

**Резюме.** З метою розробки алгоритму індивідуального прогнозування ступеня тяжкості стану у хворих на ХОЗЛ із супутньою ГХ на основі даних біохімічного дослідження сироватки крові було обстежено 100 пацієнтів – 55 з них мали поєднану патологію ХОЗЛ з ГХ, у 45 був ізольований перебіг ХОЗЛ та контрольна група складалась з 45 умовно здорових осіб. На підставі використання кластерного (деревовидного) алгоритму об'єднання об'єктів було виявлено з 100 пацієнтів 2 однорідні групи — 2 кластера. Методами дивізійного аналізу середніх показників кластерів (к-середніх) підтверджено виділення саме двох кластерів, визначено якісний склад кожного кластера, тобто кожен хворий співвіднесений до відповідного кластеру. Розподіл хворих за двома кластерами відповідає виділенню двох груп пацієнтів з різним ступенем тяжкості їх стану при поєднаному перебігу ХОЗЛ і ГХ: 1-й кластер — це група хворих з більш тяжким ступенем стану; 2-й кластер — група хворих із середнім ступенем тяжкості стану. Методами дискримінантного аналізу розроблено алгоритм індивідуального прогнозу ступеня тяжкості стану хворого з поєднаним перебігом ХОЗЛ і ГХ. Класифікація тяжкості стану будь-якого хворого з ХОЗЛ проводиться на підставі встановлених розрахункових класифікаційних функцій Y1 кластер і Y2 кластер для даного хворого. Хворий відноситься до того кластеру (групі тяжкості стану), для якого чисельне значення класифікаційної функції максимальне. Матриця класифікації за двома кластерами підтверджує її високу точність.

**Ключові слова:** хронічне обструктивне захворювання легень, гіпертонічна хвороба, математико-статистичний аналіз.

УДК [616.24-007.272-036.12+616.12-008.331.1]-018.74-008.6-07:57.083.3

**МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ ПРОГНОЗИРОВАНИЯ СТЕПЕНИ ТЯЖЕСТИ СОСТОЯНИЯ У БОЛЬНЫХ ХРОНИЧЕСКИМ ОБСТРУКТИВНЫМ ЗАБОЛЕВАНИЕМ ЛЕГКИХ С СОПУТСТВУЮЩЕЙ ГИПЕРТОНИЧЕСКОЙ БОЛЕЗНЬЮ**  
**Капустник В. А., Истомина О. В., Зайцева О. В.**

**Резюме.** С целью разработки алгоритма индивидуального прогнозирования степени тяжести состояния у больных ХОЗЛ с сопутствующей ГБ на основе данных биохимического исследования сыворотки крови было обследовано 100 пациентов — 55 из них имели сочетанную патологию ХОЗЛ с ГБ, 45 были с изолированным течением ХОЗЛ и контрольная группа состояла из 45 условно здоровых лиц. На основе использования кластерного (древовидного) алгоритма объединения объектов было выявлено из 100 пациентов 2 однородные группы — 2 кластера. Методами дивизионного анализа средних показателей кластеров (к-средних) подтверждено выделение именно двух кластеров, определены качественный состав каждого кластера, то есть каждый больной соотношен к соответствующему кластеру. Распределение больных по двум кластерам соответствует выделению двух групп пациентов с разной степенью тяжести их состояния при совместном течении ХОЗЛ и ГБ: 1-й кластер — это группа больных с более тяжелой степенью состояния; 2-й кластер — группа больных со средней степенью тяжести состояния. Методами дискриминантного анализа разработан алгоритм индивидуального прогнозирования степени тяжести состояния больного с сочетанным течением ХОЗЛ и ГБ. Классификация тяжести состояния любого больного с ХОЗЛ проводилась на основании установленных расчетных классификационных функций Y1 кластер и Y2 кластер для данного больного. Больной относится к тому кластеру (группе тяжести состояния), для которого численное значение классификационной функции максимально. Матрица классификации по двум кластерам подтверждает ее высокую точность.

**Ключевые слова:** хроническое обструктивное заболевание легких, гипертоническая болезнь, математико-статистический анализ.

UDC [616.24-007.272-036.12+616.12-008.331.1]-018.74-008.6-07:57.083.3

### MATHEMATICS AND STATISTICAL ANALYSIS OF THE PREDICTION OF THE SEVERITY OF THE CONDITION OF PATIENTS WITH CHRONIC OBSTRUCTIVE PULMONARY DISEASE IN COMBINATION WITH HYPERTENSION

Капустник В. А., Истомина О. В., Зайтсева О. В.

**Abstract.** Chronic obstructive pulmonary disease (COPD) in combination with hypertension is one of the most common comorbid conditions in internal medicine clinic. Until now, there has been no comprehensive study of the functional state of the endothelium in patients with COPD in combination with hypertension and there is no clear algorithm for predicting the severity of these patients.

*The purpose* of this stage of work was the development of an algorithm for individual prediction of the severity of the condition in patients with COPD concomitant with hypertension based on biochemical serum.

Were examined 145 patients which were divided into 3 groups: the main group included patients with COPD in combination with hypertension (55 persons); comparison group consisted of patients with isolated COPD (45 people); group control made relatively healthy patients (45 people). All the subjects underwent the following studies: clinical and biochemical blood tests, levels of catecholamines, metabolites and synthase of nitric oxide, S-nitrosothiols, VEGF, level of biogenic elements, vasoactive hormones, peptides and neurotransmitters were determined.

Based on the use of the cluster (tree) algorithm of combining objects, 2 homogeneous groups — 2 clusters — were identified from 100 patients. By methods of divisive analysis of average cluster indices (k-means), it was confirmed that two clusters were isolated, the qualitative composition of each cluster was determined, that is, each patient was correlated to the corresponding cluster. The distribution of patients in two clusters corresponds to the allocation of two groups of patients with different degrees of severity of their condition in the joint course of COPD and hypertension: 1<sup>st</sup> cluster is a group of patients with a more severe degree of condition; 2<sup>nd</sup> cluster — a group of patients with an average severity of the condition. Methods of discriminant analysis have developed an algorithm for individual prediction of the degree of severity of a patient with a combined course of COPD and hypertension. The classification of the severity of the condition of any patient with COPD was performed based on the established classification functions of the Y1 cluster and Y2 cluster for the given patient. The patient refers to the cluster (the severity group of the state) for which the numerical value of the classification function is maximal. The classification matrix for two clusters confirms its high accuracy.

**Keywords:** chronic obstructive pulmonary disease, hypertension, mathematics and statistical analysis.

Рецензент — проф. Катеренчук І. П.

Стаття надійшла 10.03.2017 року