

УДК [577.175.47+ 616.414]: 004.891.2

## РАЗРАБОТКА МЕТОДИКИ ОПРЕДЕЛЕНИЯ РЕФЕРЕНТНЫХ ИНТЕРВАЛОВ РАЗЛИЧНЫХ ПОКАЗАТЕЛЕЙ ПРИ РАЗВИТИИ КОМПЕНСАТОРНЫХ РЕАКЦИЙ В ОРГАНИЗМЕ



**О.С. Спиридонова**  
аналитик данных  
ООО "HiQo-Solutions"



**Н.В. Карлович**  
врач-эндокринолог, кандидат  
медицинских наук

ООО "HiQo-Solutions"

Белорусский государственный медицинский университет, Республика Беларусь

E-mail: [olyasprdnv@gmail.com](mailto:olyasprdnv@gmail.com)

**О.С. Спиридонова**

*Работает в HiQo-Solutions в должности системного аналитика.*

**Н.В. Карлович**

*Врач-эндокринолог, высшая квалификационная категория, кандидат медицинских наук. Является членом Белорусской ассоциации врачей, Белорусской ассоциации эндокринологов, БОМО «Эндокринология и метаболизм», Европейской ассоциации эндокринологов.*

**Аннотация.** Определение оптимальных значений параметров, уровни которых начинают резко возрастать при наличии различных патологий, и развивающихся вследствие этого компенсаторных механизмов является актуальной задачей на сегодняшний день. В работе описывается методика для определения референтных интервалов различных показателей при развитии компенсаторных реакций в организме. Основными этапами данной методики являются понижение размерности и кластеризация данных. Кластеризация пациентов осуществляется не на исходных значениях параметров, а на данных, свернутых автоэнкодером, что позволяет кластеризовать пациентов по наиболее ключевым комбинациям признаков. Алгоритм данной методики может быть применен в различных областях для более качественного контроля медикаментозной терапии при лечении различных патологий, сопровождающихся компенсаторными изменениями показателей.

**Ключевые слова:** кластеризация, понижение размерности, автоэнкодер, компенсаторные механизмы.

### **Введение.**

При развитии патологических процессов в организме, для поддержания гомеостаза начинают действовать различные компенсаторные механизмы. Результатом таких процессов является повышение или понижение некоторых параметров до уровней, сильно отличающихся от нормы, что именно и обеспечивает нормальное функционирование организма в условиях патологии. Так, например, повышение уровня гемоглобина у курящих лиц, является следствием снижения сатурации кислорода, в связи с изменением легочной ткани под воздействием никотина и смол. Повышение гемоглобина выше нормы, способствует удержанию в норме сатурации у курящих лиц [1]. Важным моментом является знание граничных значений параметров, при которых уже необходимо назначать коррекцию. На сегодняшний момент, определение оптимальных уровней данных параметров при развитии компенсаций, является актуальной задачей, что позволит не

пропустить момент начала коррекции. Цель исследования – разработать методику с применением нейросетевых алгоритмов, для определения референтных интервалов различных показателей при развитии компенсаторных реакций в организме.

#### **Материалы и методы.**

Для разработки и апробации методики были взяты 2 исследования. Первое исследование в области эндокринологии, для определения референтных интервалов паратгормона (ПТГ) у пациентов с различными стадиями хронической болезни почек (ХБП). При наличии ХБП, для поддержания в норме параметров костного обмена, развивается компенсаторное повышение уровня паратгормона. Вследствие этого, у пациентов с ХБП становится проблематичным выявление истинного повышения данного гормона, при котором требуется медикаментозная коррекция. Анализ данных осуществляли с помощью пакета прикладных программ STATISTICA 10,0 (Stat Soft, 2001) и Anaconda (Python), изучение статистических параметров распределения признаков – с использованием методов описательной статистики. Для сравнения групп между собой применяли соответствующие параметрические и непараметрические тесты. Проверку соответствия реального распределения переменных нормальному осуществляли с помощью теста Колмогорова–Смирнова. Критический уровень значимости при доверительном оценивании параметров распределения принимали равным 0,05. Первоначальная задача при разработки методики определения референтного интервала паратгормона заключалась в том, чтобы разбить (кластеризовать) пациентов на группы в зависимости от состояния костного обмена и выделить группу с наиболее оптимальными показателями. Первоначально, для понижения размерности применяли нейросетевой алгоритм (автоэнкодер), состоящий из энкодера и декодера. Входной слой состоял из 21 нейрона. Энкодер состоял из двух слоев: 63 нейрона в первом слое и 42 нейрона во втором. Декодер имел соответствующую структуру, но слои располагались в обратном порядке. Скрытый слой состоял из 3 нейронов. Данное ограничение на количество нейронов в скрытом слое позволяло модели отбирать наиболее значимые параметры (комбинации параметров) и более качественно проводить кластеризацию. Далее на модель подавали обучающие данные. Сигнал снимался со скрытого слоя и обрабатывался с помощью методов кластеризации без учителя (метод k-средних, DBSCAN). Далее в сформированных группах и подгруппах был проведен анализ выживаемости пациентов, методом Каплана–Мейера. В соответствии с результатами кластеризации и анализа выживаемости устанавливали группу с оптимальным состоянием костного метаболизма. Значения уровня ПТГ в данной группе принимали за референтный интервал [2]. Второе исследование, проводилось в области кардиологии, для определения референтных значений параметров микроциркуляции бульбарной конъюнктивы глаза у пациентов с сердечно-сосудистыми заболеваниями [3]. Методика определения референтных интервалов была в целом идентична первому исследованию. Кластеризация проводилась на параметрах транспорта кислорода у пациентов с сердечно-сосудистыми заболеваниями. После выделения группы пациентов с наиболее оптимальными параметрами транспорта кислорода, определялись референтные значения параметров микроциркуляции бульбарной конъюнктивы глаза.

#### **Результаты.**

Одним из основных моментов являлось то, что кластеризация производилась не на исходных значениях параметров костного обмена, а на данных, свернутых автоэнкодером, что позволило кластеризовать пациентов по наиболее ключевым комбинациям признаков. На рисунке 1 приведены результаты кластеризации методом k-средних, параметров костного обмена после понижения размерности, полученных со скрытого слоя автоэнкодера.

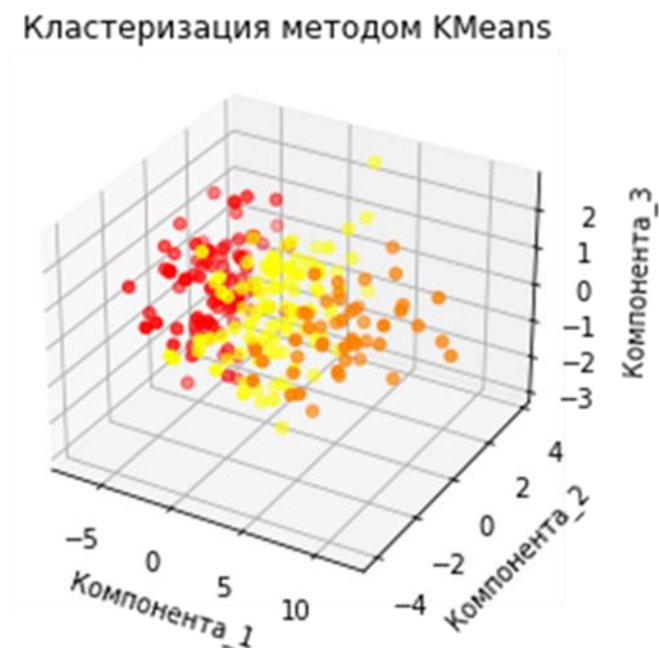


Рисунок 1. Кластеризация пациентов методом K-means на скрытом слое автоэнкодера

Функция минимизации на тестовых данных при понижении размерности (средне-квадратичная ошибка) после обучения автоэнкодера составила 0.038. Качество кластеризации и количество кластеров определялось с помощью метода Элбо. По результатам кластеризации выделяли группу с наиболее оптимальными параметрами костного обмена и смотрели распределение уровня паратгормона в этой группе. На рисунке 2 приведено распределение уровня ПТГ в группе с оптимальными значениями показателей костного обмена. За верхнюю границу был выбран 90% квантиль, он составил 506 пг/мл.

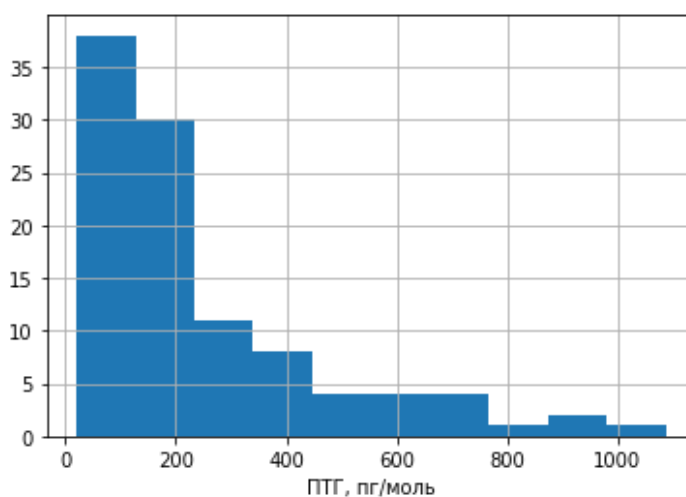


Рисунок 2. Гистограмма распределения ПТГ в группе с оптимальными показателями костного обмена

### Заклучение.

Разработанная методика на основе алгоритмов машинного обучения позволит определять референтные интервалы различных показателей при развитии компенсаторных

реакций в организме. Основой данной методики является кластеризация, проведенная на данных после уменьшения размерности. Кластеризация может быть сделана с помощью различных методов (метод k-средних, DBSCAN, иерархическая кластеризация). Понижение размерности также можно осуществлять различными способами (автоэнкодер на основе полностью связанной сети прямого распространения, метод главных компонент). Алгоритм данной методики может быть применен в различных областях для более качественного контроля медикаментозной терапии при лечении различных патологий, сопровождающихся компенсаторными изменениями показателей.

#### ***Список использованных источников***

[1] Correlation between anemia and smoking: Study of patients visiting different outpatient departments of Integral Institute of Medical Science and Research, Lucknow /Waseem SMA, Alvi AB// Natl J Physiol Pharm Pharmacol 2020;10(02):149-154

[2] Определение референтного интервала паратгормона у пациентов с различными стадиями хронической болезни почек / Н. В. Карлович [и др.] // Вест. Нац. акад. наук Беларуси. Сер. мед. наук. – 2021. – Т. 18, № 2. – С. 000–000. <https://doi.org/10.29235/1814-6023-2021-18-2-000-000>

[3] Microvascular dysfunction in patients with incomplete myocardial revascularization in the long-term period of primary percutaneous coronary intervention (diagnosis and treatment opportunities)/Tsapaeva N.L., Tarashkevich N.V., Mironova E.V., Jn Shojan, Konstantinova E.E., Chernoglaz P.F., Yurlevich D.I.// Neotlozhnaya kardiologiya i kardioovaskulyarnye riski [Emergency cardiology and cardiovascular risks] 1, 126–133.

## **DEVELOPMENT OF A METHOD FOR DETERMINING REFERENCE INTERVALS OF DIFFERENT INDICATORS DURING THE DEVELOPMENT OF COMPENSATORY REACTIONS IN THE ORGANISM**

***O.S. SPIRIDONOVA***  
*Systems Analyst "HiQo-Solutions"*

***N.V. KARLOVICH,***  
*Ph. D. (Med.), Endocrinologist*

*"HiQo-Solutions"*  
*Belarusian State Medical University, Republic of Belarus*  
*E-mail: olyasprdnv@gmail.com*

**Abstract.** Determining the optimal values of parameters, the levels of which begin to increase sharply in the presence of various pathologies, and the compensatory mechanisms developing as a result of this, is an urgent task today. The paper describes a technique for determining the reference intervals of various indicators in the development of compensatory reactions in the body. The main steps of this technique are dimensionality reduction and data clustering. Clustering of patients is carried out not on the initial values of the parameters, but on data convolved by an autoencoder, which allows clustering patients according to the most key combinations of features. The algorithm of this technique can be applied in various areas for better control of drug therapy in the treatment of various pathologies, accompanied by compensatory changes in indicators.

**Keywords:** clustering, dimensionality reduction, autoencoder, compensatory mechanisms.