

УДК 51-76

DOI: [10.26102/2310-6018/2023.41.2.011](https://doi.org/10.26102/2310-6018/2023.41.2.011)

## Кластеризация пациентов на основе их функциональных, клинических и антропометрических показателей для построения моделей оценки биовозраста

О.В. Лимановская<sup>1,2</sup>✉, И.В. Гаврилов<sup>1,2</sup>, В.Н. Мещанинов<sup>1,2</sup>

<sup>1</sup>Уральский государственный медицинский университет Министерства здравоохранения Российской Федерации, Екатеринбург, Российская Федерация

<sup>2</sup>Центр специализированных видов медицинской помощи «Институт медицинских клеточных технологий», Екатеринбург, Российская Федерация  
[limanovskaya@mail.ru](mailto:limanovskaya@mail.ru)✉

**Резюме.** Кластерный анализ получил широкое распространение как инструмент анализа медицинских данных для выделения групп пациентов. Но несмотря на широкое применение кластерного анализа довольно редко встречаются работы, где наравне с выделением групп пациентов, математически обосновываются признаки, по которым произошло разделение на группы. Для решения этой задачи может быть применен метод, получивший название кластеризация с учителем, суть которого заключается в применении методов многоклассовой классификации с использованием меток кластеров как целевой переменной. В данной работе этот метод применен для выделения показателей, по которым произойдет разделение групп пациентов базы данных медицинской организации ГАУЗ СО «СОКП Госпиталь для ветеранов войн» и ГАУЗ СО «Институт медицинских клеточных технологий» за 1995-2022 гг. в объеме 6440 записей данных пациентов. В качестве метода кластеризации использован метод HDBscan, в качестве метода верификации полученных кластеров пациентов использован метод CatBoost в режиме многоклассовой классификации. В результате получено 4 кластера, разделившиеся по гендерному признаку и по состоянию пациента. Для выявления статистических отличий полученных кластеров проведен АВ анализ данных кластеров с использованием критерия Крускала-Уолиса. Результаты АВ анализа показали, что полученные кластеры имеют статистически значимые отличия по всем включенным в анализ функциональным параметрам. Далее был проведен АВ анализ различий функциональных показателей пациентов в амбулаторном и стационарном статусе для женского и мужского кластера. Для АВ анализа использовался перестановочный критерий и бутстрап с построением доверительных интервалов средних из генерированных в бутстрапе выборок.

**Ключевые слова:** кластеризация с учителем, АВ анализ групп, геропротективное воздействие, прогнозирование эффективности лечения, биовозраст.

**Благодарности:** работа выполнена частично в рамках и при поддержке государственного задания Федерального государственного бюджетного образовательного учреждения Уральский государственный медицинский университет Министерства здравоохранения Российской Федерации, тема: "Индивидуализация подбора комплексной геропротективной терапии", номер 121030900298-9.

**Для цитирования:** Лимановская О.В., Гаврилов И.В., Мещанинов В.Н. Кластеризация пациентов на основе их функциональных, клинических и антропометрических показателей для построения моделей оценки биовозраста. *Моделирование, оптимизация и информационные технологии*. 2023;11(2). URL: <https://moitvvt.ru/ru/journal/pdf?id=1335> DOI: 10.26102/2310-6018/2023.41.2.011

## Clustering of patients based on their functional, clinical and anthropometric indicators for the construction of models for assessing bio-age

O.V. Limanovskaya<sup>1,2</sup>✉, I.V. Gavrilov<sup>1,2</sup>, V.N. Meshchaninov<sup>1,2</sup>,

<sup>1</sup>Ural State Medical University of the Ministry of Health of the Russian Federation,  
Ekaterinburg, the Russian Federation

<sup>2</sup>Specialized Medical Care Center of Medical Cell Technology Institute, Ekaterinburg,  
the Russian Federation  
[limanovskaya@mail.ru](mailto:limanovskaya@mail.ru)✉

**Abstract.** Cluster analysis has become a widely used tool for analyzing medical data to identify groups of patients. But despite the widespread use of cluster analysis, it is rare to find publications where the identification of groups of patients and the attributes by which the division into groups occurred are mathematically justified. To solve this problem, a method called clustering with a teacher can be applied, the essence of which is to apply multiclass classification methods using cluster labels as a target variable. In this paper, this method is employed to identify indicators by which groups of patients will be divided in the databases of the autonomous public health care institutions SOCP Hospital for War Veterans and Institute of Medical Cell Technologies for years 1995-2022 in volume 6440. The HDBscan method was used for clustering method, and the CatBoost method in the multiclass classification mode was used as a verification method for the obtained clusters of patients. As a result, 4 clusters were obtained divided by gender and the patient's condition. In order to identify statistical differences between the obtained clusters, an AB analysis of these clusters was carried out by means of the Kruskal-Walis criterion. The results of the AB analysis showed that the clusters have statistically significant differences in all functional parameters included in the analysis. Further, an AB analysis of the differences in the functional indicators of patients in outpatient and inpatient status for the female and male cluster was carried out. For the AB analysis, a permutation criterion and a bootstrap were used with the construction of confidence intervals of averages from samples generated in the bootstrap.

**Keywords:** supervision clustering, AB analysis, geroprophylactic treatment, prediction of treatment effectiveness, bio-growth.

**Acknowledgments:** the research was partially carried out within the framework and with the support of the state task of the Federal State Budgetary Educational University of the Ministry of Health of the Russian Federation, subject: "Individualization of complex geroprophylactic therapy selection", study number 121030900298-9.

**For citation:** Limanovskaya O.V., Gavrilov I.V., Meshchaninov V.N. Clustering of patients based on their functional, clinical and anthropometric indicators for the construction of models for assessing bio-age. *Modeling, Optimization and Information Technology*. 2023;11(2). URL: <https://moitvvt.ru/ru/journal/pdf?id=1335> DOI: 10.26102/2310-6018/2023.41.2.011 (In Russ.).

### Введение

Кластерный анализ нашел широкое применение в обработке медицинских данных в качестве инструмента выделения групп пациентов [1-7], имеющих схожие характеристики. Выделение групп пациентов часто проводится среди пациентов, имеющих определенное заболевание [1, 2, 5-8], и гораздо реже для здоровых людей [9]. Качество полученных кластеров и их характеристики оцениваются медицинскими специалистами с клинической точки зрения с использованием стандартных статистических методов, таких как оценка средних значений параметров и сравнительный анализ групп. Такой подход к оценке результатов кластеризации не позволяет понять по каким именно параметрам произошло выделение групп пациентов

из выборки, что затрудняет понимание медицинской значимости результатов кластеризации. Наличие информации о том, по каким именно параметрам выделились группы пациентов, позволит сильно расширить сферу применения кластерного анализа в медицинских исследованиях. Для решения этой задачи используется смешанный подход, получивший название кластеризация с учителем [10], сочетающий в себе как метод обучения без учителя (кластеризацию), так и метод обучения с учителем – классификацию. Суть этого метода состоит в том, что полученные после кластеризации метки кластеров данных используются как целевая переменная для задачи классификации. В этом подходе сначала проводится кластеризация данных, а на следующем этапе классификация этих же данных по числу классов, равных числу кластеров. В качестве целевой переменной в задаче классификации служат метки кластеров, полученных на первом этапе. При таком подходе появляется возможность получить значения важностей параметров, по которым произошла классификация. На основе важностей параметров можно выделить наиболее значимые параметры в каждом полученном кластере и, следовательно, выявить параметры, по которым разделились кластеры. Кроме того, для каждого кластера при использовании этого подхода можно получить оценку качества, пользуясь метриками классификации –  $f1$  мера, точность, полнота. Таким образом, метод кластеризации с учителем является перспективным методом для выявления клинических показателей, по которым идет разделение пациентов на группы.

Данная работа посвящена применению метода кластеризации с учителем для выделения групп пациентов на основе их антропометрических, клинических и функциональных параметров и выявления параметров, по которым эти группы разделились. Полученные группы пациентов послужат основой для построения моделей биовозраста в геронтологии для каждой группы и станут входными данными для разрабатываемой авторами агентной модели старения.

### Материалы и методы

Данные. Для выявления когорт пациентов использована база данных медицинской организации ГАУЗ СО «СОКП Госпиталь для ветеранов войн» и ГАУЗ СО «Институт медицинских клеточных технологий» (г. Екатеринбург) за 1995-2022 гг. в объеме 6440 записей данных пациентов в возрасте от 15 до 93 лет (средний возраст 51 год), содержащей 2525 пациентов женского пола, 3915 пациентов мужского пола. Анализировались 13 показателей из предоставленного набора функциональных параметров:

- 1) АДС – артериальное давление систолическое в мм рт. ст., среднее значение 128 мм рт. ст.;
- 2) АДД – артериальное давление диастолическое в мм рт. ст., среднее значение 79 мм рт. ст.;
- 3) АДП – разность между систолическим и диастолическим давлением в мм рт. ст., среднее значение 49 мм рт. ст.;
- 4) ЗДВдох – время задержки дыхания на вдохе в секундах, среднее значение 45 с;
- 5) ЗДВыдох – время задержки дыхания на выдохе в секундах, среднее значение 27 с;
- 6) ЖЕЛ – жизненная емкость легких в мл, среднее значение 3481 мл;
- 7) масса – масса тела в кг, среднее значение 78 кг;
- 8) аккомодация хрусталика глаза в диоптриях, среднее значение 4,8 диоптрий;
- 9) острота слуха в бел, среднее значение 31 бел;
- 10) статическая балансировка в секундах, среднее значение 17 с;

11) пол, 2525 женщин, 3915 мужчин;

12) статус пациента – амбулаторный или стационарный, 4379 стационарных пациентов, 2061 амбулаторный пациент.

Метод кластеризации. Кластеризация данных с целью выявления когорт проводилась методом HDBSCAN [11]. Метод основан на определении плотности объектов и объединении соседних объектов выборки в кластеры. Ближайшие к объекту выборки соседи, расстояние до которых считается на основе параметров объектов выборки по заданной метрике, объединяются в кластер. Объекты, не имеющие соседей, считаются выбросами. В данной работе для расчета расстояния использовалась Евклидова метрика. Размерность минимального кластера задавалась равной 50.

Верификация результатов кластеризации. Для оценки результатов кластеризации проведена многоклассовая классификация методом CatBoost [12], где в качестве меток класса для каждого объекта выборки использовались метки кластера. CatBoost основан на использовании ансамбля классификаторов, что дает более высокую точность модели классификации, чем использование одного классификатора. Алгоритм разработан Яндексом. Алгоритм основан на использовании деревьев решений. При обучении модели идет последовательное построение деревьев решений таким образом, что каждое следующее дерево решений обучается на ошибках предыдущего. Такой подход позволяет получить более быстрый градиентный спуск по плоскости решений. Использование классификатора позволило получить оценку каждого F1 мерой, которая дает усредненную оценку модели по полноте и точности. Кроме того, применение классификатора позволило получить оценку важности факторов модели, используя метод классификатора feature\_importance.

## Результаты

Подготовка данных. Из базы данных были выбраны пациенты, имеющие записи по функциональным параметрам. Дублирующие записи по одному пациенту были удалены. В итоге получена база данных из 4242 пациента от 15 до 93 лет, 2458 мужчин и 1784 женщин. Для приведения данных к единому масштабу (от 0 до 1) использовалось минимаксное преобразование данных по следующим формулам (1,2):

$$X_{\text{сред}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}) \quad (1)$$

$$X_{\text{преобразованное}} = X_{\text{сред}} * (\text{max} - \text{min}) + \text{min}. \quad (2)$$

Категориальные данные (пол и состояние) были приведены к бинарному виду.

Кластеризация. Обработанные данные были кластеризованы методом HDBScan, и получены 4 кластера (Рисунок 1).

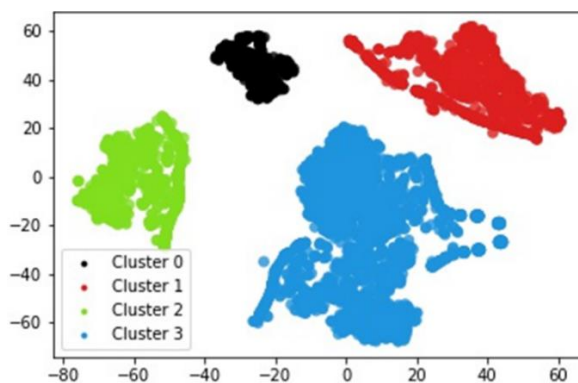


Рисунок 1 – Результаты кластеризации  
Figure 1 – Clustering results

Проведен анализ полученных кластеров (Таблица 1), в ходе которого дана оценка кластера в виде F1-меры, рассчитан гендерный состав кластера и число пациентов с различным состоянием (легкое состояние – амбулаторные пациенты и среднее – стационарные пациенты).

Таблица 1 – Оценка результатов кластеризации  
Table 1 – Validation of clustering results

Номер кластера	F1-мера	Объем кластера, кол-во пациентов	Число мужчин в кластере	Число женщин в кластере	Число амбулаторных пациентов в кластере	Число стационарных пациентов в кластере
0	1	344	344	0	344	0
1	1	991	0	991	991	0
2	1	793	0	793	0	793
3	1	2114	2114	0	0	2114

По данным Таблицы 1, полученные кластеры можно описать следующим образом. Кластер 0 – амбулаторные пациенты мужского пола. Объем кластера 344 пациента. Кластер 1 – амбулаторные пациенты женского пола. Объем кластера 944 пациента. Кластер 2 – стационарные пациенты женского пола. Объем кластера 793 пациентки. Кластер 3 – стационарные пациенты мужского пола. Объем кластера 2114 пациента. Самый большой кластер – это кластер стационарных пациентов мужского пола, самый маленький кластер – это амбулаторные пациенты мужского пола.

Все кластеры хорошего качества, о чем свидетельствует значение F1 меры, равное 1 для всех кластеров. Кластеры разделились по гендерному признаку и типу пациентов, что также подтверждается таблицей важности параметров кластеров, представленной в Таблице 2, из которой видно, что для каждого кластера именно пол и состояние пациента являются признаками, по которому выделился кластер из выборки.

Таблица 2 – Важности параметров, по которым выделились кластеры  
Table 2 – Importance of clustering attributes

Параметр	Кластер 0	Кластер 1	Кластер 2	Кластер 3
состояние	52,83	62,19	41,18	31,75
пол	41,68	33,24	55,82	64,76
возраст	0,55	0,26	0,17	0,20
АДС	0,27	0,26	0,17	0,20
АДД	0,36	0,33	0,22	0,25
АДП	0,71	0,21	0,14	0,16
ЗДВдох	0,43	0,46	0,30	0,35
ЗДВыдох	0,39	0,64	0,43	0,50
ЖЕЛ	1,10	0,48	0,31	0,37
Масса тела	0,38	0,24	0,16	0,18
Аккомодация	0,42	0,99	0,65	0,76
Острота слуха	0,43	0,34	0,23	0,26
Стат. балансировка	0,44	0,33	0,22	0,25

Анализ полученных кластеров. Для каждого кластера рассчитаны для всех показателей средние значения и 95 % доверительный интервал по формуле (1). Результаты приведены в Таблице 3.

Таблица 3 – Средние значения параметров кластеров  
Table 3 – Mean values of clustering attributes

Параметр	Кластер 0	Кластер 1	Кластер 2	Кластер 3
возраст	45,83 ± 1,62	47,04 ± 0,95	56,15 ± 1,06	51,12 ± 0,73
АДС	129,62 ± 1,45	124,76 ± 0,87	130,30 ± 1,28	129,50 ± 0,66
АДД	81,72 ± 1,02	78,81 ± 0,55	79,43 ± 0,68	79,71 ± 0,42
АДП	48,35 ± 1,16	46,12 ± 0,60	51,60 ± 0,95	49,89 ± 0,53
ЗДВдох	53,78 ± 1,78	42,68 ± 0,74	38,94 ± 0,83	45,82 ± 0,77
ЗДВвыдох	29,40 ± 0,86	26,52 ± 0,50	24,64 ± 0,44	26,58 ± 0,38
ЖЕЛ	3703,13 ± 81,47	3246,72 ± 29,96	3020,79 ± 50,54	3622,09 ± 40,61
Масса тела	80,49 ± 1,31	71,72 ± 0,79	74,22 ± 0,87	79,19 ± 0,53
Аккомодация	5,36 ± 0,25	5,21 ± 0,12	4,81 ± 0,13	4,82 ± 0,10
Острота слуха	27,06 ± 1,46	25,86 ± 0,56	28,18 ± 0,78	33,46 ± 0,80
Стат. балансировка	22,59 ± 2,19	18,83 ± 1,07	12,63 ± 1,09	16,81 ± 0,87

Как видно из Таблицы 3, доверительные интервалы по всем используемым в кластеризации параметрам не пересекаются, что позволяет предположить, что функциональные показатели пациентов в кластерах имеют статистически значимые отличия.

Сравнительный анализ выделенных групп пациентов. Для выявления статистических отличий полученных кластеров по функциональным показателям использовался критерий Крускала-Уолиса [13] при нулевой гипотезе о равенстве средних. Проведенные расчеты показали, что значения  $p\_value$  близки к нулю, что позволяет отвергнуть нулевую гипотезу о равенстве средних значений исследуемых показателей и подтверждает наличие статистически значимой разницы между кластерами по всем параметрам, используемым при кластеризации.

### Обсуждение

Результаты кластеризации показали наличие 4 кластеров, имеющих статистически важные отличия по всем функциональным показателям пациентов, входящих в них. С клинической точки зрения важно различие между клиническими и стационарными пациентами одного пола, то есть как изменяются функциональные показатели пациента при изменении его состояния с амбулаторного пациента на стационарного. Нами ранее в работе [13] выдвинуто предположение о наличии гендерной разницы в механизмах старения, что дает возможность предположить наличие гендерной разницы в изменениях функциональных показателей при изменении состояния пациента. Поэтому представляет интерес сравнить не только 4 кластера между собой, но и сравнить изменения функциональных параметров в разных состояниях для мужчин и женщин отдельно. Предполагается, что разница между функциональными параметрами амбулаторных и стационарных пациентов женского пола будет статистически значимо отличаться от таковой у пациентов мужского пола. Поэтому для каждого кластера данных проведено сравнение различий усредненных функциональных параметров амбулаторных и стационарных пациентов мужского и женского пола. Результаты приведены в Таблице 4.

Таблица 4 – Разницы средних значений функциональных показателей между амбулаторными и стационарными пациентами мужского и женского пола

Table 4 – Differences of functional parameter means between outpatient and inpatient male and female patients

	Муж	Жен
АДС	0,12	5,56
АДД	2,01	0,63
АДП	1,54	5,51
ЗДВдох	7,96	3,75
ЗДВывдох	2,82	1,88
ЖЕЛ	81,04	226,45
Масса тела	1,30	2,54
Аккомодация	0,54	0,40
Острота слуха	0,40	2,32
Стат.балансировка	5,78	6,45

Как видно из Таблицы 4, наблюдаются существенные гендерные различия в изменении разницы средних показателей амбулаторных и стационарных пациентов. Таким образом, можно предположить, что есть существенное влияние гендера на разницу функциональных показателей пациентов в разном состоянии.

С целью выявления статистической значимости гендерного различия проведен статистический анализ различий функциональных показателей между амбулаторными и стационарными пациентами. Для этого рассчитывалась разница функциональных параметров между амбулаторными и стационарными пациентами и анализировались полученные векторы данных.

Сначала проведена проверка нормальности распределения полученных разностей с помощью критерия Шапиро-Уилка [14]. Результаты показали, что все данные имеют не нормальное распределение. Поэтому для оценки статистической разницы между мужской и женской группой использовались 95 % доверительные интервалы, построенные на средних значениях выборок, полученных из изначальных выборок, методом бутстрапа и перестановочный критерий [15]. Полученные значения p-value перестановочного критерия для разницы всех функциональных параметров близки к нулю. Выводы по статистической значимости различий делались на основе пересечения доверительных интервалов и величин p-value используемого критерия. Результаты приведены в Таблице 5.

Таблица 5 – 95 % доверительные интервалы разницы функциональных показателей и биовозраста и значения p\_value перестановочного критерия

Table 5 – 95 % confidence intervals of the difference between functional indicators and bio-age and the p\_value values of the permutation criterion

	Жен	Муж	Является ли различие значимым
АДС	-6,74..-4,12	0,18..0,84	да
АДД	-1,59..-0,03	2,61..3,90	да
АДП	-6,06..-4,20	-2,27..-1,71	да
ЗДВдох	2,98..4,72	8,39..9,28	да
ЗДВывдох	1,35..2,42	3,03..3,41	да
ЖЕЛ	184,04..273,84	129,07..170,06	да

Таблица 5 (продолжение)  
Table 5 (extended)

	Жен	Муж	Является ли различие значимым
Масса тела	-3,59..-1,55	1,32..1,95	да
Аккомодация	0,26..0,57	0,43..0,54	сомнительно
Острота слуха	-3,13..-1,58	-6,51..-5,75	да
Стат. балансировка	4,96..7,61	4,31..5,29	сомнительно

Как видно из Таблицы 5, практически все функциональные параметры, за исключением аккомодации и статической балансировки, имеют статистически значимые гендерные различия.

Результаты анализа различий данных кластеров позволили подтвердить гипотезу о наличии гендерного влияния на разницу функциональных показателей между состояниями пациентов. Поскольку модель биовозраста, предложенная нами в работе [16], основана на функциональных данных пациентов, возникает необходимость построения моделей прогнозирования биовозраста отдельно для мужчин и для женщин.

### Заключение

В работе проведен кластерный анализ функциональных данных пациентов, который показал наличие 4 кластеров в выборке, разделенных по гендерному признаку и состоянию пациента. Далее была рассчитана разница функциональных показателей амбулаторных и стационарных пациентов. Полученные значения усреднены и проведено сравнение значений между мужской и женской группами. Статистический анализ гендерного различия показал, что для большинства функциональных показателей оно статистически значимо. Поскольку предложенная нами ранее модель определения биовозраста [16] основана на функциональных данных пациентов, то целесообразно построение отдельных моделей для определения биовозраста для мужчин и для женщин. Полученные в результате кластеризации группы пациентов будут использованы для построения моделей биовозраста.

### СПИСОК ИСТОЧНИКОВ

1. Криштоп В.В., Пахрова О.А. Применение кластерного и корреляционного анализа для оценки гемореологических показателей у больных эссенциальной артериальной гипертензией. *Успехи современного естествознания*. 2014;9:11–16.
2. Deckersbach T., Peters A.T., Sylvia L.G., Gold A.K., da Silva Magalhaes P.V., Henry D.B., Frank E., Otto M.W., Berk M., Dougherty D.D., Nierenberg A.A., Miklowitz D.J. A Cluster Analytic Approach to Identifying Predictors and Moderators of Psychosocial Treatment for Bipolar Depression: Results from STEP-BD. *J Affect Disord*. 2016;203:152–157. DOI: 10.1016/j.jad.2016.03.064.
3. O'Regan A., Hannigan A., Glynn L., Garcia Bengoechea E., Donnelly A., Hayes G., Murphy A.W., Clifford A.M., Gallagher S., Woods C.B. A cluster analysis of device-measured physical activity behaviours and the association with chronic conditions, multi-morbidity and healthcare utilisation in adults aged 45 years and older. *Woods Preventive Medicine Reports*. 2021;24:101641–101651. DOI: 10.1016/j.pmedr.2021.101641.
4. Serpa Neto A., Bos L.D., Campos P.P.Z.A., Hemmes S.N.T., Bluth T., Calfee C.S., Ferner M., Güldner A., Hollmann M.W., India I., Kiss T., Laufenberg-Feldmann R., Sprung J., Sulemanji D., Unzueta C., Vidal Melo M.F., Weingarten T.N., Tuip-de Boer



- A.M., Pelosi P., Gama de Abreu M., Schultz M.J. Association between pre-operative biological phenotypes and postoperative pulmonary complications an unbiased cluster analysis. *Eur J Anaesthesiol.* 2018;35:702–709. DOI: 10.1097/EJA.0000000000000846.
5. Gagnon P., Casaburi R., Saey D., Porszasz J., Provencher S., Milot J., Bourbeau J., O'Donnell D.E., Maltais F. Cluster Analysis in Patients with GOLD 1 Chronic Obstructive Pulmonary Disease *PLoS ONE.* 2015;10(4):e0123626. DOI: 10.1371/journal.pone.0123626.
  6. Sharma A., Zheng Y., Ezekowitz J.A., Westerhout C.M., Udell J.A., Goodman S.G., Armstrong P.W., Buse J.B., Green J.B., Josse R.G., Kaufman K.D., McGuire D.K., Ambrosio G., Chuang L.M., Lopes R.D., Peterson E.D., Holman R.R. Cluster Analysis of Cardiovascular Phenotypes in Patients With Type 2 Diabetes and Established Atherosclerotic Cardiovascular Disease: A Potential Approach to Precision Medicine. *Diabetes Care.* 2022;45:204–212. DOI: 10.2337/dc20-2806.
  7. Демченко М.В., Каширина И.Л., Фирюлина М.А. Кластеризация состояний пациентов для модели назначения схем лечения атеросклероза. *Вестник ВГУ, серия: системный анализ и информационные технологии.* 2021;2:126–137. DOI: <https://doi.org/10.17308/sait.2021.2/3509>.
  8. Konno S., Taniguchi N., Makita H., Nakamaru Yu., Shimizu K., Shijubo N., Fuke S., Takeyabu K., Oguri M., Kimura H., Maeda Yu., Suzuki M., Nagai K., Yo. M. Ito, Wenzel S.E., Nishimuka M. Distinct phenotypes of Cigarette Smokers Identified by Cluster Analysis of Patients with Severe Asthma. *AnnalsATS.* 2015;12(12):1771–1780. DOI: 10.1513/AnnalsATS.201507-407OC.
  9. O'Regan A., Hannigan A., Glynn L., Garcia Bengoechea E., Donnelly A., Hayes G., Murphy A.W., Clifford A.M., Gallagher S., Woods C.B. A cluster analysis of device-measured physical activity behaviours and the association with chronic conditions, multi-morbidity and healthcare utilisation in adults aged 45 years and older. *Preventive Medicine Reports.* 2021;24:101641. DOI: 10.1016/j.pmedr.2021.101641.
  10. Al-Harbi S.H., Rayward-Smith V.J. Adapting k-means for supervised clustering. *Journal of Applied Intelligence.* 2006;24(3):219–226.
  11. McInnes L., Healy J. Accelerated Hierarchical Density Based Clustering. *IEEE International Conference on Data Mining Workshops (ICDMW), IEEE.* 2017;1:33–42.
  12. CatBoost. Доступно по: <https://catboost.ai> (дата обращения 21.01.2023).
  13. Мещанинов В.Н., Гаврилов И.В., Мякотных В.С., Щербаков Д.Л. Гендерные демографические и статистические особенности старения человека. *Новые информационные технологии в образовании и науке.* 2022;2(6):65–72.
  14. Кобзарь А.И. Прикладная математическая статистика. М.: Физматлит; 2006. 466–468 с.
  15. Колядин В.Л. Пермутационные критерии как универсальный непараметрический подход к проверке статистических гипотез. *Радиоэлектроника и информатика.* 2002;3(20):7–14.
  16. Лимановская О.В. Гаврилов И.В. Мещанинов В.Н. Щербаков Д.Л. Колос Е.Н. Моделирование биологического возраста пациентов на основе их функциональных показателей. *Моделирование, оптимизация и информационные технологии.* 2021;9(2). Доступно по: <https://moitvvt.ru/ru/journal/pdf?id=966> DOI: 10.26102/2310-6018/2021.33.2.028

## REFERENCES

1. Krishtop V.V., Pakhrova O.A. Application of cluster and correlation analysis to assess hemorheological parameters in patients with essential arterial hypertension. *Uspekhi sovremennogo estestvoznaniya = Advances in current natural sciences*. 2014;9:11–16. (In Russ.).
2. Deckersbach T., Peters A.T., Sylvia L.G., Gold A.K., da Silva Magalhaes P.V., Henry D.B., Frank E., Otto M.W., Berk M., Dougherty D.D., Nierenberg A.A., Miklowitz D.J. A Cluster Analytic Approach to Identifying Predictors and Moderators of Psychosocial Treatment for Bipolar Depression: Results from STEP-BD. *J Affect Disord*. 2016;203:152–157. DOI: 10.1016/j.jad.2016.03.064.
3. O'Regan A., Hannigan A., Glynn L., Garcia Bengoechea E., Donnelly A., Hayes G., Murphy A.W., Clifford A.M., Gallagher S., Woods C.B. A cluster analysis of device-measured physical activity behaviours and the association with chronic conditions, multi-morbidity and healthcare utilisation in adults aged 45 years and older. *Woods Preventive Medicine Reports*. 2021;24:101641–101651. DOI: 10.1016/j.pmedr.2021.101641.
4. Serpa Neto A., Bos L.D., Campos P.P.Z.A., Hemmes S.N.T., Bluth T., Calfee C.S., Ferner M., Güldner A., Hollmann M.W., India I., Kiss T., Laufenberg-Feldmann R., Sprung J., Sulemanji D., Unzueta C., Vidal Melo M.F., Weingarten T.N., Tuip-de Boer A.M., Pelosi P., Gama de Abreu M., Schultz M.J. Association between pre-operative biological phenotypes and postoperative pulmonary complications an unbiased cluster analysis. *Eur J Anaesthesiol*. 2018;35:702–709. DOI: 10.1097/EJA.0000000000000846.
5. Gagnon P., Casaburi R., Saey D., Porszasz J., Provencher S., Milot J., Bourbeau J., O'Donnell D.E., Maltais F. Cluster Analysis in Patients with GOLD 1 Chronic Obstructive Pulmonary Disease *PLoS ONE*. 2015;10(4):e0123626. DOI: 10.1371/journal.pone.0123626.
6. Sharma A., Zheng Y., Ezekowitz J.A., Westerhout C.M., Udell J.A., Goodman S.G., Armstrong P.W., Buse J.B., Green J.B., Josse R.G., Kaufman K.D., McGuire D.K., Ambrosio G., Chuang L.M., Lopes R.D., Peterson E.D., Holman R.R. Cluster Analysis of Cardiovascular Phenotypes in Patients With Type 2 Diabetes and Established Atherosclerotic Cardiovascular Disease: A Potential Approach to Precision Medicine. *Diabetes Care*. 2022;45:204–212. DOI: 10.2337/dc20-2806.
7. Demchenko M.V., Kashirina I.L., Firyulina M.A. Cluster analysis of patients' states performed in order to develop treatment strategies for patients with atherosclerosis. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Sistemnyj analiz i informacionnye tehnologii = Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies*. 2021;2:126–137. DOI: 10.17308/sait.2021.2/3509. (In Russ.).
8. Konno S., Taniguchi N., Makita H., Nakamaru Yu., Shimizu K., Shijubo N., Fuke S., Takeyabu K., Oguri M., Kimura H., Maeda Yu., Suzuki M., Nagai K., Yo. M. Ito, Wenzel S.E., Nishimuka M. Distinct phenotypes of Cigarette Smokers Identified by Cluster Analysis of Patients with Severe Asthma. *AnnalsATS*. 2015;12(12):1771–1780. DOI: 10.1513/AnnalsATS.201507-407OC.
9. O'Regan A., Hannigan A., Glynn L., Garcia Bengoechea E., Donnelly A., Hayes G., Murphy A.W., Clifford A.M., Gallagher S., Woods C.B. A cluster analysis of device-measured physical activity behaviours and the association with chronic conditions, multi-morbidity and healthcare utilisation in adults aged 45 years and older. *Preventive Medicine Reports*. 2021;24:101641. DOI: 10.1016/j.pmedr.2021.101641.

10. Al-Harbi S. H., Rayward-Smith V.J. Adapting k-means for supervised clustering. *Journal of Applied Intelligence*. 2006;24(3):219–226.
11. McInnes L., Healy J. Accelerated Hierarchical Density Based Clustering. *IEEE International Conference on Data Mining Workshops (ICDMW), IEEE*. 2017;1:33–42.
12. CatBoost. Available from: <https://catboost.ai> (accessed on 21.01.2023).
13. Meshchaninov V.N., Gavrilov I.V., Myakotny`kh V.S., Shherbakov D.L. Genderny`e demograficheskie i statisticheskie osobennosti stareniya cheloveka. *Novy`e informacziorny`e tekhnologii v obrazovanii i nauke*. 2022;2(6):65–72. (In Russ.).
14. Kobzar A.I. Applied mathematical statistics. M.: Fizmatlit; 2006. 466–468 p. (In Russ.).
15. Kolyadin V.L. Permutation criteria as a universal nonparametric approach to testing statistical hypotheses. *Радиоэлектроника и информатика = «Radioelectronics & Informatics» Journal*. 2002;3(20):7–14. (In Russ.).
16. Limanovskaya O.V. Gavrilov I.V. Meshchaninov V.N. Shcherbakov D.L. Kolos E.N. Modeling the biological age of the patients based on their functional indicators. *Modeling, Optimization and Information Technology*. 2021;9(2). Available from: <https://moitvvt.ru/ru/journal/pdf?id=966> DOI: 10.26102/2310-6018/2021.33.2.028 (In Russ.).

#### ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

**Лимановская Оксана Викторовна**, кандидат химических наук, старший научный сотрудник лаборатории антивозрастных технологий «Институт медицинских клеточных технологий», старший научный сотрудник отдела общей патологии Уральского государственного медицинского университета Минздрава РФ. Екатеринбург, Российская Федерация.

*e-mail:* [limanovskaya@mail.ru](mailto:limanovskaya@mail.ru)  
ORCID: [0000-0002-2084-3916](https://orcid.org/0000-0002-2084-3916)

**Oksana Viktorovna Limanovskaya**, Candidate of Chemical Sciences, Senior Researcher at the Laboratory of Anti-Aging Technologies of Specialized Medical Care Center of Medical Cell Technology Institute, Senior Researcher at the Department of Common Patology of Ural State Medical University of the Ministry of Health of the Russian Federation, Yekaterinburg, the Russian Federation.

**Гаврилов Илья Валерьевич**, кандидат биологических наук, доцент кафедры биохимии Уральского государственного медицинского университета Минздрава РФ», старший научный сотрудник лаборатории антивозрастных технологий «Институт медицинских клеточных технологий», Екатеринбург, Российская Федерация

*e-mail:* [iliagavrilov18@yandex.ru](mailto:iliagavrilov18@yandex.ru)  
ORCID: [0000-0003-0806-1177](https://orcid.org/0000-0003-0806-1177)

**Iliya Valeriyavich Gavrilov**, Candidate of Biological Sciences, Associate Professor at the Department of Biochemistry of Ural State Medical University of the Ministry of Health of the Russian Federation, Senior Researcher at the Laboratory of Anti-Aging Technologies of Specialized Medical Care Center of Medical Cell Technology Institute, Yekaterinburg, the Russian Federation.

**Мещанинов Виктор Николаевич**, доктор медицинских наук, профессор, заведующий кафедрой биохимии Уральского государственного медицинского университета Минздрава РФ, заведующий лабораторией антивозрастных технологий «Институт медицинских клеточных технологий», Екатеринбург, Российская Федерация

*e-mail:* [mv-02@yandex.ru](mailto:mv-02@yandex.ru)  
ORCID: [0000-0001-7928-2503](https://orcid.org/0000-0001-7928-2503)

**Viktor Nikolaevich Meshchaninov**, Doctor of Medical Sciences, Professor, Head of the Department of Biochemistry of Ural State Medical University of the Ministry of Health of the Russian Federation, Head of the Laboratory of Anti-Aging Technologies of Specialized Medical Care Center of Medical Cell Technology Institute, Yekaterinburg, the Russian Federation.

*Статья поступила в редакцию 20.03.2023; одобрена после рецензирования 24.04.2023;  
принята к публикации 26.05.2023.*

*The article was submitted 20.03.2023; approved after reviewing 24.04.2023;  
accepted for publication 26.05.2023.*