

УДК: 004.89 004.94 612.13

М. В. ВОЙТИКОВА¹, Р. В. ХУРСА²

**ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ
ДЛЯ ДИАГНОСТИКИ ЛАТЕНТНЫХ НАРУШЕНИЙ ГЕМОДИНАМИКИ:
УСЛОВИЯ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ**

¹*Институт физики им. Б. И. Степанова Национальной академии наук Беларуси, Минск, Беларусь,
e-mail: voitikova@imaph.bas-net.by*

²*Белорусский государственный медицинский университет, Минск, Беларусь,
e-mail: rvkhursa@tut.by*

Исследовались особенности применения построенной с помощью интеллектуального анализа данных номограммы для классификации гемодинамических состояний по величинам артериального давления (АД) при суточном мониторинговании и случайном измерении АД стандартными измерительными приборами. При нанесении на номограмму вектора характерных особенностей АД, координаты которого есть набор коэффициентов линейной регрессии систолического, диастолического и пульсового давления, гемодинамическое состояние диагностируется по расположению вектора в определенных границах номограммы. Показано, что для практического использования номограммы потребуется получение минимально 10 успешных (исключающих артефакты) измерений АД при малой вариабельности его показателей или 25–30 измерений при высокой.

Ключевые слова: интеллектуальный анализ данных, диагностическая номограмма, гемодинамика.

M. V. VOITIKOVA¹, R. V. KHURSA²

**DIAGNOSTICS OF LATENT HEMODYNAMICS BY DATA MINING:
CONDITIONS OF PRACTICAL USE**

¹*B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus, Minsk, Belarus,
e-mail: voitikova@imaph.bas-net.by*

²*Belarusian State Medical University, Minsk, Belarus, e-mail: rvkhursa@tut.by*

We proposed a diagnostic nomogram (map) for hemodynamic state classification, based on linear regression modeling of blood pressure (BP) parameters and Data Mining algorithm. Mapping the vector of the BP characteristics, which coordinates are the coefficients of linear regression of the systolic, diastolic on pulse pressure, the pathological changes of hemodynamics inherent in hypertension, prehypertension or clinically latent hemodynamic disturbances can be diagnosed. The normal physiology of cardiovascular interaction corresponds to the vector location within certain limits on the nomogram. We have investigated the factors influencing the accuracy of the diagnostics in a clinical setting.

Keywords: Data Mining, diagnostic nomogram, hemodynamics.

Введение. Алгоритмы интеллектуального анализа данных Data Mining в последнее десятилетие широко используются для поддержки принятия решений в различных сферах человеческой деятельности [1], в том числе в медицинской диагностике [2]. Технологии Data Mining основаны на поиске в массиве данных скрытых закономерностей или шаблонов информации, действующих на основе правил, формализующих экспертные знания, что выгодно отличает их от статистических методов анализа, ориентированных на проверку заранее сформулированных гипотез. Так, при помощи обученной на примерах машины опорных векторов, использующей вектор характерных особенностей сигнала артериального давления (АД), представленного набором коэффициентов линейной регрессии систолического, диастолического и пульсового давления, авторами была

разработана диагностическая номограмма для моделирования процессов в гемодинамике – науке, изучающей механизмы движения крови в сердечно-сосудистой системе.

Известно, что установлению клинически явной артериальной гипертензии (АГ) предшествуют латентные нарушения сердечно-сосудистой системы, например повышенная вариабельность АД, увеличение сосудистой жесткости, изменение параметров центральной гемодинамики и др. [3]. Однако диагностика таких отклонений требует дополнительных инструментальных методов, использующихся с данной целью только по обоснованным показаниям.

Целью настоящего исследования является определение условий практического применения оригинальной номограммы гемодинамических состояний для диагностики латентных нарушений кровообращения пациентов при суточном мониторинге артериального давления (СМАД) [4, 5] и случайных измерениях АД стандартными измерительными приборами.

Диагностическая номограмма (карта) гемодинамических состояний создавалась с помощью обученной на примерах машины опорных векторов (алгоритм интеллектуального анализа данных Data Mining). Она представляет собой систему координат коэффициентов линейной регрессии систолического давления по пульсовому давлению, полученных при СМАД (отдельно за дневной, ночной и суточный периоды), на которую наносятся параметры регрессии тестируемого пациента. Нами были определены границы параметров регрессии, соответствующие нормотензивному гармоническому кровообращению пациента, и критерии, указывающие на патологическое изменение сердечно-сосудистой системы, свойственное АГ, вторичной гипотензии или нормотензии с дисфункциональными гемодинамическими нарушениями, включая клинически латентные у практически здоровых людей. Созданная диагностическая номограмма позволяет выделить 9 гемодинамических классов: артериальная гипотензия или артериальная гипертензия гармонического либо дисфункциональных типов (диастолического, систолического); гармонический или дисфункциональные (диастолический, систолический) типы у нормотензивных лиц, а также квази-гипертензия (Квази-АГ) [4–7]. Среди латентных гемодинамических нарушений в когорте нормотензивных пациентов особое внимание привлекает класс, названный нами Квази-АГ, который можно рассматривать как предгипертензию, поскольку по ряду параметров СМАД пациенты этого класса сходны с пациентами АГ до начала лечения. Диагностика данного состояния, переходного к АГ, методами математического анализа временных рядов АД с применением номограммы представляется важным в клинической практике.

Регрессионное моделирование гемодинамики с применением интеллектуального анализа данных. Применение диагностической номограммы предполагает регрессионное моделирование гемодинамики по результатам измерений СМАД. Величина АД является интегральной характеристикой сердечно-сосудистой системы, зависит от объема циркулирующей крови, согласованной работы сердца и сосудов, реологических свойств крови и периферического сопротивления сосудов. Артериальное давление характеризуется величинами систолического давления S (максимальное давление при сокращении сердечной мышцы), диастолического давления D (минимальное давление между ударами сердца в момент расслабления сердечной мышцы), а также пульсового давления W (по определению: $W = S - D$). Пульсовое давление W является барическим эквивалентом пульсового объема крови, включает в себе результат взаимодействия сократительной функции сердца, растяжимости артерий и величины волны отражения.

Линейное регрессионное моделирование параметров АД с пульсовым давлением W в качестве аргумента использовалось нами для диагностики типа гемодинамики и раннего определения гемодинамических нарушений. Коэффициенты получаемой регрессии характеризуют индивидуальное кровообращение как процесс сердечно-сосудистого взаимодействия в продвижении крови [8]. Регрессионная модель кровообращения в аналитическом виде определяется линейным уравнением $S = Q + aW$, где a и Q – коэффициенты регрессии. Коэффициент Q определяет величину давления в области затухающей пульсовой волны, а значения коэффициента a косвенно задает соотношение систолической и диастолической составляющих процесса кровообращения.

Применение алгоритмов интеллектуального анализа данных [9] к параметрам измерения АД позволило получить новую значимую информацию для классификации гемодинамики пациента.

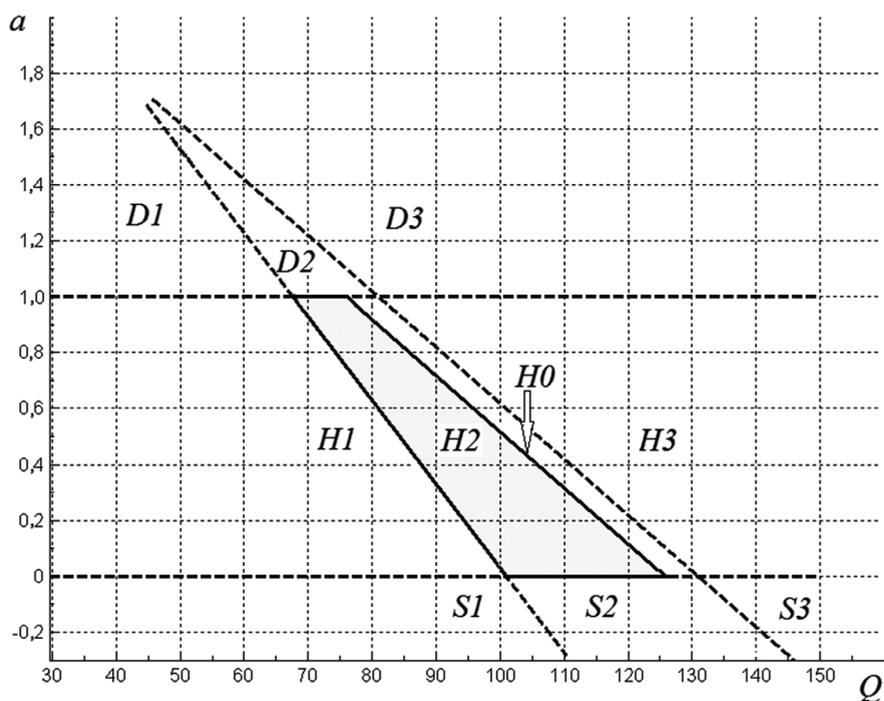


Рис. 1. Номограмма гемодинамических состояний (дневной период СМАД): область D – диастолический дисфункциональный тип кровообращения; область S – систолический дисфункциональный тип; область H – гармонический тип (в том числе $H1$ – вторичная гипотензия; $H2$ – нормотензия; $H0$ – Квази-АГ; $H3$ – АГ)

Алгоритм Support Vector Machine (SVM), или машина опорных векторов, обученный на библиотеке примеров АД пациентов с различной гемодинамикой (гипертензией, нормотензией и эпизодами вторичной гипотензии), был положен в основу диагностической номограммы (рис. 1), на которую наносится 2-мерный вектор характерных признаков (характеристический вектор) $\{Q, a\}$ для диагностики гемодинамики тестируемого пациента [4, 5]. Гемодинамические классы разделяются прямыми: $a = -0,03Q + 3,03$; $a = -0,02Q + 2,52$; $a = -0,02Q + 2,62$; $a = 0$; $a = 1$. Гемодинамический класс пациентов с параметрами в области $H0$ (см. рис. 1) – Квази-АГ – был выделен классификатором на основе SVM в той части номограммы, где множество векторов АД нормо- и гипертензивных лиц перекрываются в зоне гармонического типа кровообращения ($0 < a < 1$). Это означает, что гемодинамика лиц, попавших в эту область, характеризуется гармоническим кровообращением, но при этом сходна с гемодинамикой пациентов с АГ (т. е. Квази-АГ обозначает область, где равновероятно представлены регрессионные параметры АД нормо- и гипертензивных пациентов). Было показано, что эта подгруппа представляет пациентов с переходным к АГ гемодинамическим состоянием, которое можно рассматривать как прегипертензию у пациентов, имеющих на момент исследования нормальные показатели АД [5, 6].

Для правильной диагностики гемодинамических состояний по предлагаемой классификации, особенно Квази-АГ (учитывая «узкий» участок области характеристических векторов на номограмме, задающийся неравенством $-0,02Q + 2,52 < a < -0,02Q + 2,62$ при $0 < a < 1$), задача достоверного определения коэффициентов регрессии $\{Q, a\}$ является актуальной.

Факторы, влияющие на результаты регрессионного анализа артериального давления.

Основными факторами, способными негативно влиять на достоверность определения коэффициентов регрессии $\{Q, a\}$, могут являться недостаточная продолжительность ряда измерений АД (число точек L , по которым строится регрессия) и наличие артефактов измерений (несоблюдение правил измерения, недостаточная квалификация измеряющего, техническая неисправность прибора и др.). Поэтому в первую очередь ставилась задача определения минимального числа успешных измерений АД при СМАД, достаточного для достоверного определения коэффициентов регрессии. Для анализа использовались 43 временных ряда АД различных пациентов длиной от 1000 до 10 000 измерений (база данных работы [10]) и собственные данные –

58 образцов СМАД, выполненных аппаратными средствами серии *BPLab* (длительностью не менее 20 ч). Ряды разной длины были получены при «прореживании» исходного ряда измерений СМАД для достижения рядов нужной длины (минимально 6 измерений АД). На рис. 2 и 3 приведены вычисленные коэффициенты регрессии для разной длины ряда величин АД пациентов *c2_a40551* (из базы данных [10]) и *K40T02* (собственные данные). Как видно, значительные изменения величин регрессионных коэффициентов $\{Q, a\}$ возникают при уменьшении длины ряда менее 15 измерений АД (для наглядности зависимостей $\{Q, a\}$ от длины ряда L на рис. 2–5 использован логарифмический масштаб). Следует отметить, что для рядов СМАД с малым отклонением от кривой суточного ритма допустимо использовать для получения регрессии ряды длиной 6–10 измерений. Отклонения АД от среднего уровня выражаются стандартным отклонением от среднего (SD) или производными от SD индексами [11]. Пример полученных регрессионных коэффициентов для ряда АД с малым отклонением от кривой суточного ритма ($SD < 10$ мм. рт. ст. для дневных измерений, пациент *K2M01*) показан на рис. 4.

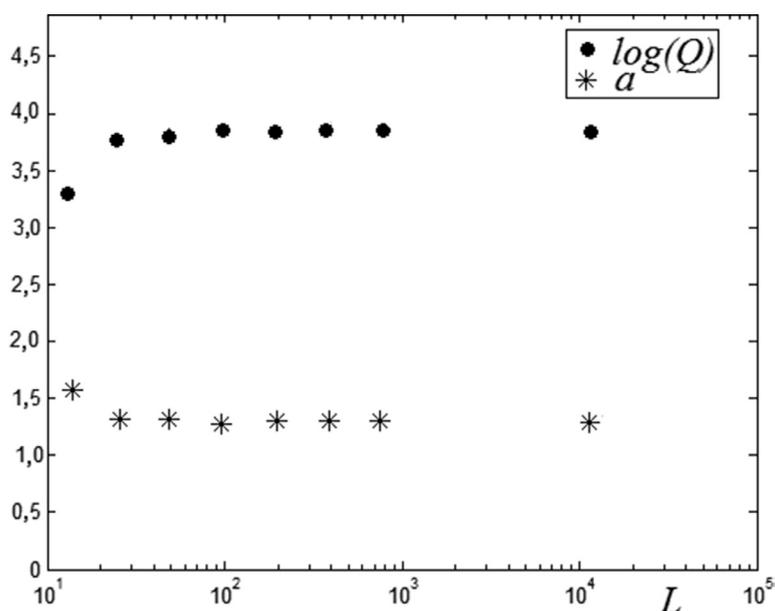


Рис. 2. Зависимость коэффициентов линейной регрессии $\log(Q)$, a от числа точек измерений L (пациент *c2_a40551*)

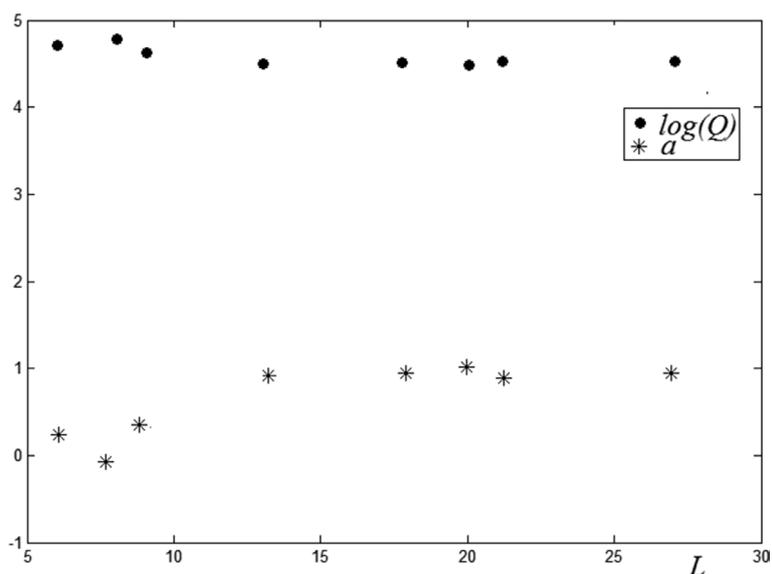


Рис. 3. Зависимость $\log(Q)$, a от длины ряда L (пациент *K40T02*)

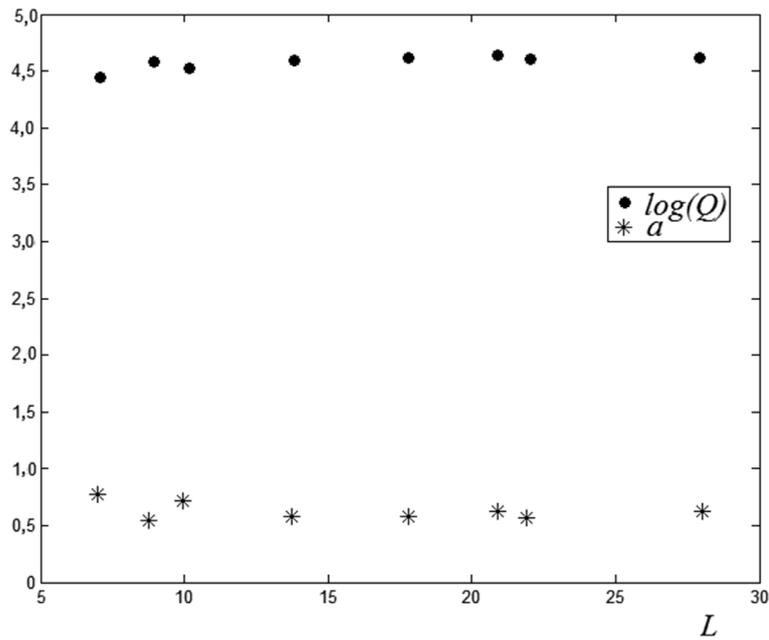


Рис. 4. Зависимость $\log(Q)$, a от длины ряда L (пациент K2M01)

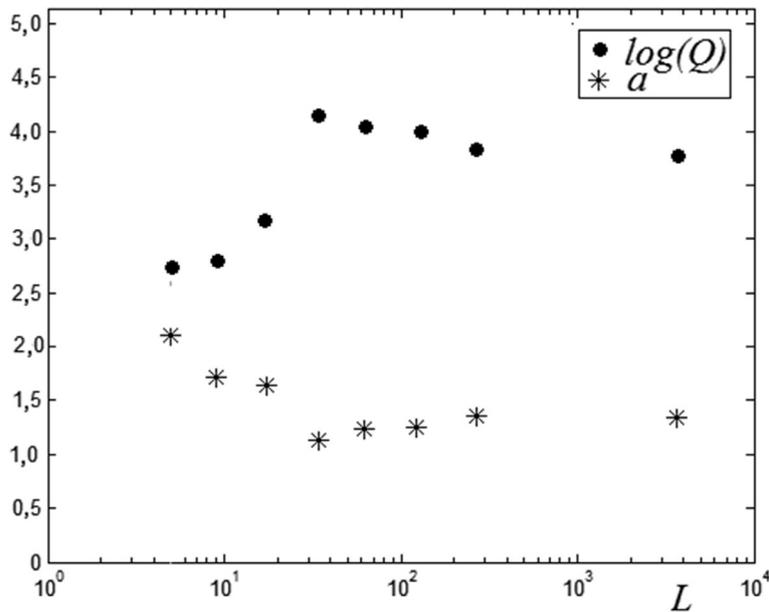


Рис. 5. Зависимость $\log(Q)$, a от длины ряда L (пациент c2_a40473, ряд АД которого содержит артефакт)

Таким образом, можно утверждать, что малая длина ряда АД (менее 15–20 измерений) становится препятствием применения номограммы гемодинамических состояний, если ряд СМАД имеет большие отклонения от кривой суточного ритма ($SD > 15$ мм. рт. ст. для дневных измерений). Средняя длина ряда величин АД при суточном мониторинговании с интервалом измерений 15–20 мин составляет 30–50 за дневной период, поэтому указанные ограничения не влияют на достоверность диагностики АД при 25 и более успешных измерений.

Рассмотрим также другую возможную причину получения недостоверных результатов при регрессионном моделировании АД. Как правило, при СМАД аппаратными средствами автоматически удаляются грубые артефакты (ошибки измерения, возникающие при двигательной активности пациента, нарушениях сердечного ритма, отсутствии данных АД и др.). В противном случае артефакты приводят к неправильной оценке результатов СМАД, как видно из рис. 5, где

анализировался содержащий кратковременный артефакт (отсутствие сигнала) ряд АД (пациент c2_a40473, база данных [10]). Действие артефакта нивелируется только при длительности ряда АД, превышающего 100 измерений. Таким образом, для качественного математического анализа временных рядов АД требуется выбраковка медицинским экспертом или аппаратными средствами участков ряда, содержащих артефакты.

Применение номограммы для данных артериального давления при ручном мониторинге. Наши исследования показали возможность применения номограммы при регулярных измерениях АД, выполненных пациентами самостоятельно с помощью ручных тонометров в течение периода времени от 2 недель до нескольких месяцев. Было установлено, что для диагностики гемодинамики допустимо использовать ряд АД с достаточным числом измерений для получения минимально шести-семи уровней пульсового давления W в терминологии работы [8] при точности измерения давления 5 мм. рт. ст. Оказалось, что при низкой вариабельности короткие ряды

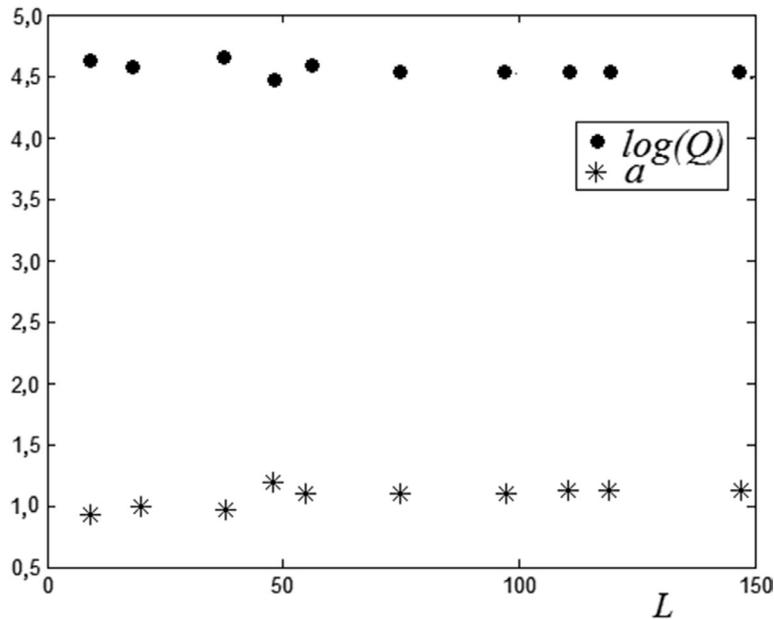


Рис. 6. Зависимость $\log(Q)$, a от длины ряда L (пациент X)

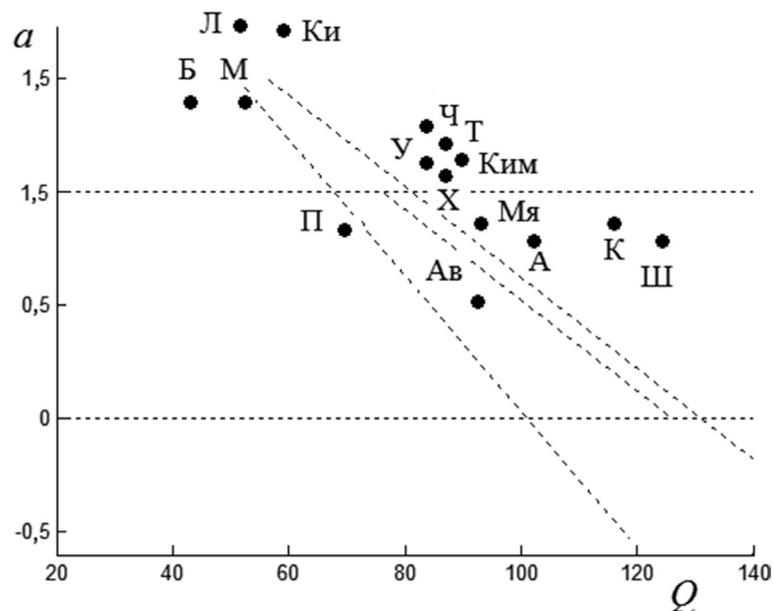


Рис. 7. Положение параметров АД пациентов при измерении давления ручными тонометрами на номограмме $\{Q, a\}$

АД (6–10 измерений) являются репрезентативными, как и длинные ряды (40–150 измерений). На рис. 6 представлена зависимость коэффициентов линейной регрессии $\log(Q)$, a от числа точек измерений L , выполненных ручным тонометром. Однако при высокой вариабельности АД ($SD > 15$ мм. рт. ст.), как и при СМАД, для анализа необходимы 25–30 измерений.

Применение номограммы для данных АД, полученных при самостоятельном измерении АД (15 пациентов, 18 образцов АД с длиной ряда от 24 до 150 измерений), показало, что параметры пациентов с выраженной артериальной гипертензией 2–3-й степени ассоциируются с областью $H3$ на номограмме (пациенты А, К, Мя, Ш) либо с областью $D3$ (пациенты Ким, Т, У, Х), а при высокой вариабельности АД – с областью $D1$ – $D2$ (пациенты Б, М) (рис. 7). Параметры АД для нормотензивного пациента располагались в области $H2$ (пациент Ав; см. рис. 7).

Заключение. Применение номограммы дополняет и расширяет диагностические возможности СМАД в сомнительных случаях гипотензии или артериальной гипертензии, а также функциональной диагностики клинически латентных нарушений сердечно-сосудистой системы нормотензивных лиц.

В настоящей работе показано, что для практического использования оригинальной номограммы гемодинамических состояний при суточном мониторинге артериального давления и случайных измерениях АД стандартными измерительными приборами потребуется получение минимального числа успешных (исключающих артефакты) измерений АД. Минимальное число подлежащих анализу измерений при малой вариабельности показателей АД ($SD < 10$ мм. рт. ст.) для дневных измерений систолического давления S составляет 10 измерений, а при высокой вариабельности АД ($SD > 15$ мм. рт. ст.) достигает 25–30 измерений. Учитывая, что средняя длина ряда дневных измерений АД при суточном мониторинге составляет 40–50, указанные ограничения не влияют на достоверность диагностики АД при 25 и более успешных измерений.

Все вышеуказанное делает диагностическую номограмму простым и надежным инструментом обследования пациентов в клинической практике для оценки индивидуальной гемодинамики, диагностики латентных гемодинамических нарушений и ранних стадий артериальной гипертензии, а также вероятности развития гипотензивных состояний.

Список использованной литературы

1. Чубукова, И. А. Data Mining / И. А. Чубукова. – М.: Интернет-университет информационных технологий: Бином. Лаборатория знаний, 2006. – (Серия «Основы информационных технологий»).
2. Кузнецова, А. В. Возможности использования методов Data Mining при медико-лабораторных исследованиях для выявления закономерностей в массивах данных / А. В. Кузнецова, О. В. Сенько // Врач и информ. технологии. – 2005. – № 2. – С. 38–46.
3. 2013 ESH/ESC Guidelines for the management of arterial hypertension / The Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC) // J. Hypertens. – 2013. – Vol. 31 (7). – P. 1281–1357.
4. Войткова, М. В. Применение интеллектуального анализа данных для классификации гемодинамических состояний / М. В. Войткова, А. П. Войтович, Р. В. Хурса // Врач и информ. технологии. – 2013. – № 1. – С. 32–41.
5. Войткова, М. В. Номограмма гемодинамических состояний по параметрам артериального давления / М. В. Войткова, Р. В. Хурса // Технология живых систем. – 2014. – № 2. – С. 45–53.
6. Хурса, Р. В. Суточное мониторирование артериального давления с использованием интеллектуального анализа данных: новые диагностические возможности / Р. В. Хурса, М. В. Войткова // Артериальная гипертензия. – 2015. – № 2 (40). – С. 34–42.
7. Хурса, Р. В. Новые маркеры прегипертензии: функциональная диагностика гемодинамических классов с помощью интеллектуального анализа данных суточного мониторирования артериального давления / Р. В. Хурса // Артериальная гипертензия и профилактика сердечно-сосудистых заболеваний: материалы VIII Междунар. конф., 21–22 мая 2015 г., Витебск: прил. к журн. Кардиология в Беларуси. – 2015. – С. 263–266.
8. Хурса, Р. В. Пульсовое давление крови: роль в гемодинамике и прикладные возможности в функциональной диагностике / Р. В. Хурса // Мед. новости. – 2013. – № 4. – С. 13–19.
9. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP: учеб. пособие / А. А. Барсегян [и др.]. – СПб.: БХВ-Петербург, 2007.
10. The MIMIC II Project database: <http://physionet.org/physiobank/database/mimic2db>
11. Горбунов, В. М. Современные представления о вариабельности артериального давления / В. М. Горбунов // Рациональная фармакотерапия в кардиологии. – 2012. – Т. 8 (6). – С. 810–818.

Поступила в редакцию 18.09.2015