

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

© КОЛЛЕКТИВ АВТОРОВ, 2017
УДК 614.253.8:681.518

*Колесниченко О.Ю.¹, Мазелис А.Л.², Николаев А.Э.², Мартынов А.В.³,
Пулит В.В.³, Смородин Г.Н.⁴, Колесниченко Ю.Ю.⁵*

INTRANET-ПАЦИЕНТ В МЕДИЦИНСКИХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

¹ГК «Ремедиум», 105082, г. Москва;

²ФГБОУ ВО «Владивостокский государственный университет экономики и сервиса»,
690014, г. Владивосток;

³Компания СП.АРМ, 197227, г. Санкт-Петербург;

⁴Академическое партнерство Dell EMC в России и СНГ, 199004, г. Санкт-Петербург;

⁵Бюллетень «Анализ безопасности», 125195, г. Москва

В статье рассмотрен математический кластерный анализ данных, полученных от медицинской информационной системы «qMS» за годовой период регистрации данных с трех медицинских учреждений. Для оценки стоимости, длительности лечения и объема обследований пациентов с гипертензивной болезнью сердца была специально написана компьютерная программа (автор А.Л. Мазелис) с использованием интерактивной среды Python. Кластеризация проводилась в двух направлениях: по количеству медицинских обследований и процедур (Series treatment) и по времени ожидания медицинских обследований и процедур (Series time). Были определены две группы пациентов по распределению стоимости и длительности стационарного лечения. Также по анализу данных медицинской информационной системы описан социально-медицинский портрет пациента (iПациента), страдающего гипертензивной болезнью сердца. Предложено проводить коррекцию стандартов лечения с учетом социально-медицинского портрета пациента в соответствии с обнаруженными реальными требованиями пациентов. В проанализированной выборке почти равное внимание уделяется обследованию сердца и желудочно-кишечного тракта, что свидетельствует о большой распространенности заболеваний последнего и требует повышенного диагностического внимания к ним как сопутствующим гипертонической болезни.

Ключевые слова: *медицинские информационные системы; Большие данные; кластерный анализ; гипертензивная болезнь сердца; стандарт медицинской помощи.*

Для цитирования: Колесниченко О.Ю., Мазелис А.Л., Николаев А.Э., Мартынов А.В., Пулит В.В., Смородин Г.Н., Колесниченко Ю.Ю. Intranet-пациент в медицинских информационных системах. *Социология медицины*. 2017; 16 (1): 49—55. DOI 10.18821/1728-2810-2017-16-1-49-55

Для корреспонденции: Колесниченко Ольга Юрьевна, канд. мед. наук, e-mail: oykolesnichenko@list.ru

*Kolesnichenko O.Yu.¹, Mazelis A.L.², Nikolaiev A.E.², Martynov A.V.³, Pulit V.V.³,
Smorodin G.N.⁴, Kolesnichenko Yu.Yu.⁵*

THE INTRANET-PATIENT IN MEDICAL INFORMATION SYSTEMS

¹"Remedium", 105082 Moscow, Russia;

²The Vladivostokskii state university of economics and service, 690014 Vladivostok, Russia;

³The SP.ARM company, 197227 St. Petersburg, Russia;

⁴The academic partnership Dell EMC in Russia and SIC, 199004 St. Petersburg, Russia;

⁵The bulletin "Analiz bezopasnosti", 125195 Moscow, Russia

The article considers a mathematical cluster analysis of data obtained in medical information system "qMS" for annual period of data registration in three medical institutions. To evaluate cost, duration of treatment and scope of examination of patients with hypertension heart disease a special software was developed by A.L. Mazelis using Python interactive environment. The clustering was implemented in two directions: according number of medical examinations and procedures (Series treatment) and according time of waiting for medical examinations and procedures (Series time). Two groups of patients were established according distribution of cost and duration of hospital treatment. Also, based on analysis of data of medical information system, a description of social medical portrait of patient with hypertensive heart disease was presented. A proposal was made of implementing adjustment of treatment standards considering medical social portrait of patient according to established actual demands of patients. In analyzed sampling almost equal attention is paid to examination of heart and gastrointestinal tract that testifies wide prevalence of gastrointestinal diseases that requires increased diagnostic attention to them as concomitant ones to hypertension disease.

Key words: *medical information system; big data; hypertension heart disease; medical care standard*

For citation: Kolesnichenko O.Yu., Mazelis A.L., Nikolaiev A.E., Martynov A.V., Pulit V.V., Smorodin G.N., Kolesnichenko Yu.Yu. The Intranet-patient in medical information systems. *Sociologia meditsini*. 2017; 16 (1): 49—55. DOI 10.18821/1728-2810-2017-16-1-49-55

For correspondence: Kolesnichenko O.Yu., candidate of medical sciences. e-mail: oykolesnichenko@list.ru

Conflict of interest. The authors declare absence of conflict of interest.

Acknowledgment. The study had no sponsor support.

Received 12.10.2016
Accepted 20.01.2017

Внутренняя локальная сеть медицинского лечебного учреждения (Intranet) позволяет не только собирать большие объемы данных о каждом пациенте, но и анализировать их с целью оптимизации оказания медицинской помощи и формирования портрета пациента по каждой нозологии. С точки зрения социальной значимости анализа данных медицинских информационных систем (МИС) появляется уникальная возможность гибко реагировать на общественную реакцию на реформы в сфере здравоохранения. Новое программное обеспечение, например среда программирования Python, позволяет анализировать Большие данные, полученные с МИС для построения многосложной картины лечебного процесса. Такой подход является социализированным, так как учитывает каждого отдельного пациента в математическом алгоритме.

Список шагов, которые требуются для перехода на новый технологический уровень в информатизации здравоохранения, приводит один из ведущих Data Scientist в мире Люис Фролио (Louis Frolio) [1—3]:

- ◆ Создать структуры для сбора и хранения данных с МИС.
- ◆ Организовать интероперабельные связи между врачами и специалистами в области математического моделирования и ИТ.
- ◆ Сделать данные МИС доступными для интеграции и аналитики через ИТ-системы в необходимых режимах (например, по требованию или в режиме реального времени), учредить обучающие курсы с целью подготовки команд для аналитики Больших данных МИС, продвигать информатизацию и аналитику Больших данных во всех медицинских организациях, работать над стандартизацией электронных медицинских записей (electronic health records, EHR).

Помимо функции социальной медиации в вопросах реформы здравоохранения анализ Больших данных МИС позволяет качественно влиять на основные цели здравоохранения: искать пути более эффективного лечения населения, сохранять профессиональное долголетие людей, предотвращать инвалидизацию и смертность. По результатам выполнения плана Минздрава России, информатизация медицинских учреждений достигает 70%, что соответствует уровню ведущих западных стран [4].

Материалы и методы

В рамках совместного исследовательского сотрудничества Академического Партнерства Dell EMC (Dell EMC External Research and Academic Alliances), Владивостокского государственного университета экономики и сервиса и Санкт-Петербургской ИТ-компании СП.АРМ (информационные системы в медико-социальной сфере) проводится исследование по анализу записей МИС «qMS». Проект получил название «Case-

study Big Data: iHealthCare Optimization». МИС «qMS» является уникальной «Hospital Network», объединяющей больницы в единый информационный ресурс сбора и анализа данных по принципу SaaS (Software as-a-Service) — программное обеспечение как услуга. МИС «qMS» имеет контуры Интернета, Интранета в локальных сетях, единый датацентр с ядром программного обеспечения.

Были проанализированы записи системы «qMS», сделанные в период с 1 июня 2014 г. до 15 апреля 2015 г., с трех медицинских учреждений, оказывающих стационарную медицинскую помощь. Выборка для анализа составила 685 пациентов, все они проходили стационарное лечение с диагнозом по МКБ-10 — Код: I11.9: Гипертензивная (гипертоническая) болезнь с преимущественным поражением сердца без сердечной недостаточности, или Гипертоническая болезнь II стадии. Средний возраст пациентов составил 56,5 года. Личные данные пациентов исследовательской группе не передавались, пациенты были закодированы.

ИТ-специалисты СП.АРМ перевели собранные первичные массивы Больших данных в таблицу метаданных. Закодированная таблица с метаданными подверглась разным методам анализа: в СП.АРМ сделан математический анализ с составлением уравнения линейной регрессии, во Владивостокском государственном университете экономики и сервиса проведен кластерный анализ с помощью языка программирования Python. Также был проведен статистический анализ на основе Microsoft Excel, включая определение коэффициента корреляции Пирсона.

Анализ начинался с написания технического задания (CodeBook), которое состояло из следующих разделов, обязательных для проведения аналитики Больших данных: название Case-study, главный вопрос, на который нужно искать ответ, перечень метаданных для анализа и сетка для кластерного анализа, гипотеза и предполагаемые выводы, практическая значимость, статистические методы дополнительного анализа. Сетка для кластерного анализа не является жесткой для точного следования, так как математики при моделировании могут выявить более значимые направления кластеризации. Гипотеза и предполагаемые выводы в CodeBook служат лишь ориентиром и могут меняться при получении результатов анализа.

Результаты кластерного анализа в среде Python. Для автоматизации расчетов была специально написана программа (автор А.Л. Мазелис) с использованием интерактивной среды iPython и библиотек NumPy, Pandas и Sklearn, предназначенных для анализа и обработки данных.

Кластеризация выполнена по полному набору метаданных Max Metadata (39 медуслуг), среднему набору метаданных Middle Metadata (24 медуслуги) и сокращенному набору метаданных Min Metadata (10 медуслуг).

луг) с группировкой на 2, 3 и 4 кластера (например, кластеризация полного набора метаданных на 2 кластера — Max-2, 3 кластера — Max-3, 4 кластера — Max-4), что позволило посмотреть эволюцию сегментации данных. Кластеризация проводилась по двум направлениям, фиксированным системой «qMS»: количеству проведенных медицинских обследований и процедур (Series treatment) и времени ожидания медицинских обследований и процедур (Series time). Кластеры обозначались буквой «к» и номером перед ней (например, Max-2 1k и Max-2 2k означает кластеризацию полного набора метаданных на 2 кластера 1k и 2k).

В табл. 1 представлен весь перечень полученных кластеров, которые при визуализации на графике по критерию «стоимость лечения — длительность госпитализации» (средние показатели в каждом отдельном кластере) распределились на 2 группы. Среднее арифметическое по всей выборке составляет 33 262 руб. для стоимости лечения и 9,7 сут для длительности госпитализации.

Стоит отметить, что только кластеризация способствовала выявлению данных групп, что невозможно было сделать, используя обычный статистический анализ. Также корреляционный анализ по всей выборке без кластеризации не показал сильных связей, а после кластеризации корреляционный анализ средних значений кластеров показал наличие сильной связи между длительностью госпитализации и стоимостью лечения ($r = 0,8$).

Описание кластеров по полному набору метаданных по характеристикам «стоимость лечения—длительность госпитализации». Сначала были рассмотрены кластеры по полному набору метаданных Max Metadata и их показатели по численности пациентов, средним значениям длительности госпитализации и стоимости лечения. При кластеризации по количеству проведенных медицинских обследований и процедур обнаружена группа из 427 пациентов (62,3%), лечение которых в среднем составило 28 тыс. руб. (кластер Max-4-1k Series treatment), в то время как общее среднее по всей выборке составляет 33 тыс. руб.

Также выявлена группа из 221 пациента (32,3%), лечение которых в среднем достигло 43 тыс. руб. (кластер Max-4-4k Series treatment) при средней длительности пребывания в стационаре 11,3 сут, что выше стандарта Минздрава, которым определено 10 сут (Стандарт первичной медико-санитарной помощи при первичной артериальной гипертензии). Средний возраст пациентов этого кластера составляет 56 лет.

При кластеризации по времени ожидания медицинских обследований и процедур выявлена группа из 63 пациентов (9,2%), стоимость лечения которых превысила в среднем 52 тыс. руб. (кластер Max-4 3k Series time). Из всей мультикластерной картины кластер Max-4 3k Series time имеет самую большую стоимость лечения при средней его длительности 12,8 сут.

Описание кластеров по среднему набору метаданных по характеристикам «стоимость лечения — длительность госпитализации». По среднему набору метаданных выделен кластер Middle-4 3k Series treatment с самой низкой стоимостью лечения, он очищен от обследований, которые не входят в стандарт лечения ги-

пертонической болезни. Для этого кластера стоимость лечения в среднем составила 24 939 руб. при средней длительности лечения 8 сут, численность кластера — 231 пациент (33,7%). Можно сказать, что этот кластер является желаемым ориентиром для оптимизации лечения данной категории больных (с учетом изменения цен с периода лечения).

Описание кластеров по сокращенному набору метаданных по характеристикам «стоимость лечения — длительность госпитализации» и по структуре медицинских услуг. Выявлены два выпадающих по длительности пребывания в стационаре кластера с сокращенным набором данных — Min-4 2k Series time и Min-4 3k Series treatment. Эти кластеры очищены от обширного перечня обследований до минимального количества медуслуг, которые наиболее часто оказываются данному контингенту больных с целью диагностики и лечения (электрокардиография, комбинированное суточное мониторирование ЭКГ в 3 отведениях и АД, УЗИ почек, УЗИ почечной артерии, исследование глазного дна с применением линзы Гольдмана, забор крови из пальца, забор крови из вены, инъекция внутримышечная, инъекция внутривенная, внутривенное капельное введение лекарственных средств).

Кластер Min-4 2k Series time характеризуется численностью 20 пациентов при среднем возрасте 61 год (это верхняя средняя возрастная планка среди всех кластеров) и средней стоимостью лечения 39 908 руб. при средней длительности госпитализации 15 сут. В кластер Min-4 3k Series treatment попал всего один пациент, стоимость его лечения составила 49 256 руб. при длительности госпитализации 20 сут (из-за того, что кластер содержал лишь одного пациента, данный кластер в дальнейшем детально не рассматривался).

То, что кластер Min-4 2k Series time с самым дли-

Таблица 1

Кластеры, распределенные на 2 группы по характеристикам «стоимость—длительность» стационарного лечения

Группа	Кластер
1 — 395 пациентов; стоимость лечения до 35 тыс. руб.; средняя длительность пребывания в стационаре 8,5 сут	Series treatment Max-2 1k; Middle-2 1k; Min-2 1k; Max-3 1k; Middle-3 2k; Min-3 1k; Max-4 1k; Min-4 1k
	Series time Max-2 1k; Middle-2 1k; Min-2 1k; Max-3 2k; Max-3 3k; Middle-3 1k; Middle-3 2k; Min-3 1k; Max-4 2k; Max-4 4k; Middle-4 1k; Middle-4 2k; Middle-4 3k; Min-4 3k
2 — 290 пациентов; стоимость лечения от 35 тыс. руб.; средняя длительность пребывания в стационаре 11,2 сут	Series treatment Max-2 2k; Middle-2 2k; Min-2 2k; Max-3 2k; Max-3 3k; Middle-3 1k; Middle-3 3k; Min-3 2k; Min-3 3k; Max-4 2k; Max-4 3k; Max-4 4k; Middle-4 1k; Middle-4 2k; Middle-4 4k; Min-4 2k; Min-4 4k
	Series time Max-2 2k; Middle-2 2k; Min-2 2k; Max-3 1k; Middle-3 3k; Min-3 2k; Min-3 3k; Max-4 1k; Middle-4 4k; Min-4 1k; Min-4 4k

тельным пребыванием в стационаре (в среднем 15 сут) относится к сокращенному набору данных Min, говорит о том, что задержка этих пациентов связана именно с минимальным набором медуслуг, а более конкретно — с повторным проведением ЭКГ. Также кластер Min-4 2k Series time характеризуется самым большим среди данного распределения числом анализов крови и ожиданием парентерального лечения.

В целом повторные ЭКГ-обследования проводили всего 37 (5%) пациентам. Преимущественно ЭКГ проводилась пациентам в течение первых суток от момента поступления. Можно выделить второй пул пациентов, малочисленный, который отражает повторные, отсроченные и линейно зависящие от срока госпитализации ЭКГ-обследования. В ИТ-компании СП.АРМ был проведен регрессионный анализ (В.В. Пулит) выборки метаданных всех 685 пациентов. Полученное уравнение линейной регрессии позволило сделать вывод, что на длительность пребывания пациента с гипертензивной болезнью сердца в стационаре в наибольшей степени и с высокой достоверностью оказывает влияние время ожидания ЭКГ-обследования.

Сравнение кластеров с одинаковой стоимостью лечения по структуре медуслуг на примере кластерного распределения Max-4 Series time. Рассмотрим кластерное распределение на 4 кластера по полному набору метаданных серии Series time. Кластеры Max-4 2k и Max-4 4k можно рассмотреть с позиции того, что они имеют почти одинаковые средние значения стоимости лечения. Для кластера Max-4 2k Series time стоимость лечения составила 29 866 руб. при численности кластера 234 человека и длительности лечения 7,9 сут. Для кластера Max-4 4k Series time стоимость лечения составила 29 543 руб. при численности 267 человек и длительности лечения 9,4 сут. Оба кластера укладываются в медицинский стандарт лечения данной нозологии по длительности госпитализации, имеют стоимость ниже общего среднего и сопоставимы по численности пациентов. Они отличаются тем, что пациентам из кластера Max-4 4k больше брали кровь для анализов и больше проводили внутримышечные инъекции и внутривенное вливание препаратов. При этом стоит отметить, что из всей выборки пациентов (685 человек) внутривенные капельные вливания делали 342 пациентам, из них 300 человек получили только по одному внутривенному капельному вливанью. По времени ожидания процедуры или исследования оба кластера имеют более эффективные показатели, чем два остальных.

Полный набор обследований (в дополнение к уже рассмотренному сокращенному набору) для кластерного распределения Max-4 Series time включал следующие обследования: суточное мониторирование ЭКГ в 12 отведениях, тредмил, велоэргометрия, стресс-эхоКГ, комбинированное суточное мониторирование ЭКГ в 12 отведениях, АД и реопульмограммы, спирометрия с пробой с бронхолитиком, комплексное исследование легких, триплексное исследование брахиоцефальных артерий на интра- и экстракраниальном уровне, УЗИ органов брюшной полости и почек, комплексное полисомнологическое обследование, дуплексное исследование артерий и вен нижних конечностей, дуплексное исследование вен нижних конечностей с функциональ-

ными пробами, УЗИ щитовидной железы, тест Biohit; эзофагогастроуденоскопия, колоноскопия, УЗИ молочных желез, биопсия при эндоскопии из 1 очага, УЗИ малого таза у женщин, УЗИ малого таза у мужчин, ректальное пальцевое исследование предстательной железы, взятие материала на бактериологическое, цитологическое, гормональное или ПЦР-исследование.

Сравниваемые одинаковые по стоимости кластеры Max-4 2k Series time и Max-4 4k Series time различаются тем, что кластер Max-4 2k имеет больше биопсий, а кластер Max-4 4k — больше спирометрических обследований. В обоих кластерах самое длительное ожидание относилось к полисомнологическому обследованию.

На примере сравнения схожих по стоимости лечения кластеров можно выделить две структуры (модели) медицинских услуг: в одном случае больше проводится парентеральное лечение и больше внимания уделяется обследованию кардиореспираторной системы, а в другом потребовалось проводить дополнительную диагностику сопутствующих заболеваний.

Интересен кластер Max-4 3k, так как он является самым дорогостоящим из всех, выявленных в данном исследовании. Кластер Max-4-3k Series time отличается повторным забором крови из вены, высока в нем частота внутримышечных инъекций и дольше длился курс внутривенных инъекций. Также кластер Max-4 3k Series time отличается наибольшим количеством проведенных биопсий и эзофагогастроуденоскопий, большим временем ожидания ультразвукового исследования брахиоцефальных артерий. По результатам регрессионного анализа, сделанного СП.АРМ, ультразвуковое исследование брахиоцефальных артерий входит в список факторов, в наибольшей степени и с высокой достоверностью влияющих на длительность пребывания пациентов в стационаре.

При рассмотрении общего количества всех обследований, проводимых по всей выборке, выявлено, что на первом месте по частоте оказываемых услуг находится ЭКГ, на втором УЗИ брахиоцефальных артерий, на третьем УЗИ брюшной полости. Процентное доленое распределение обследованных органов и систем следующее: исследование сердца 32% (доля от всех проведенных обследований), дыхательной системы 3%, сосудов 18%, сна 1%, эндокринной системы 11%, желудочно-кишечного тракта 24%, мочеполовой системы 11%.

Сравнение кластеров по среднему набору метаданных по структуре медуслуг на примере кластерного распределения Middle-4 Series treatment. Кластер с самой низкой стоимостью лечения Middle-4 3k Series treatment (в среднем 24 939 руб. при средней длительности лечения 8 сут) по количеству медуслуг схож с кластером Middle-4 1k (средняя стоимость лечения 36 956 руб. при длительности лечения 10 сут), но различается по времени ожидания внутривенного капельного введения лекарственных средств, этот показатель короче для кластера Middle-4 3k.

Особых отличий кластера Middle-4 3k Series treatment от других трех кластеров данного кластерного распределения по всему среднему набору обследований не выявлено, видимо, низкая стоимость лечения, сегментированная в отдельный кластер, обусловлена

Таблица 2

Распределение по характеристикам «стоимость — длительность» стационарного лечения

Группа	Показатель	Обследование													
		ЭКГ	СМЭКГ	СМЭКГ+АД	СМЭКГ+АД+РПП	Тредмил	ВЭМ	Стресс-эхоКГ	Спиро	Ис. легких	БЦА	ГД	ПСО	Сосуды н/к	УЗИ щит. ж.
1	Абсолютное число обследованных	364	63	90	144	47	4	32							
	Нормированные показатели на 100 человек по численности для каждой из групп	92,2	15,9	22,8	36,5	11,9	1	8,1							
2	Абсолютное число обследованных	281	69	55	137	33	21	71							
	Нормированные показатели на 100 человек по численности для каждой из групп	96,9	23,8	18,9	47,2	11,3	7,2	24,5							
Группа	Показатель	Обследование													
1	Абсолютное число обследованных	18	28	226	67	5	34	147	45						
	Нормированные показатели на 100 человек по численности для каждой из групп	4,6	7,1	57,2	16,9	1,3	8,6	37,2	11,4						
2	Абсолютное число обследованных	20	63	254	118	40	64	198	80						
	Нормированные показатели на 100 человек по численности для каждой из групп	6,9	21,7	87,6	40,7	13,8	22,1	68,3	27,6						
Группа	Показатель	Обследование													
1	Абсолютное число обследованных	80	30	116	30	239	17	37							
	Нормированные показатели на 100 человек по численности для каждой из групп	20,3	7,6	29,4	7,6	60,5	4,3	9,4							
2	Абсолютное число обследованных	184	85	201	111	219	8	25							
	Нормированные показатели на 100 человек по численности для каждой из групп	63,4	29,3	69,3	38,3	75,5	2,8	8,6							
Группа	Показатель	Обследование или процедура													
1	Абсолютное число обследованных	УЗИ мал/г. Ж	УЗИ мал/г. М	Предстат. ж.	ППР	в/м	в/в	в/в кап.							
	Нормированные показатели на 100 человек по численности для каждой из групп	55	84	17	23	48	19	256							
2	Абсолютное число обследованных	УЗИ мал/г. Ж	УЗИ мал/г. М	Предстат. ж.	ППР	в/м	в/в	в/в кап.							
	Нормированные показатели на 100 человек по численности для каждой из групп	13,9	21,3	4,3	5,8	12,2	4,8	64,8							
2	Абсолютное число обследованных	84	122	51	47	120	22	290							
	Нормированные показатели на 100 человек по численности для каждой из групп	28,9	42,1	17,6	16,2	41,4	7,6	100							

Примечание. ЭКГ — электрокардиография; СМЭКГ — суточное мониторирование ЭКГ в 12 отведениях; СМЭКГ + АД — суточное мониторирование ЭКГ и артериального давления; СМЭКГ + АД + РПП — комбинированное суточное мониторирование ЭКГ в 12 отведениях, АД и репульсмограммы; ВЭМ — велоэргометрия; Стресс-эхоКГ — стресс-эхокардиография; Спиро — спирометрия с пробой с бронхолитиком; Ис. легких — комплексное исследование легких; БЦА — триплексное исследование брахиоцефальных артерий на интра- и экстракраниальном уровне; ГД — исследование глазного дна с применением линзы Гольдмана; ПСО — комплексное полисомнологическое обследование; Сосуды н/к — дуплексное исследование артерий и вен нижних конечностей, дуплексное исследование вен нижних конечностей с функциональными пробами; УЗИ щит. ж. — УЗИ щитовидной железы; УЗИ мол.ж. — УЗИ молочных желез; биопсия — биопсия при эндоскопии из 1 очага; биохит — тест Vrohit; ЭГДС — эзофагогастродуоденоскопия; кол-скоп. — колоноскопия; УЗИ брюш. п. — УЗИ органов брюшной полости; УЗИ ПА — УЗИ почечной артерии; УЗИ мал/г. ж. — УЗИ малого таза у женщин; УЗИ мал/г. м — УЗИ малого таза у мужчин; предстат. ж. — ректальное пальцевое исследование предстательной железы; ПЦР — взятие материала на бактериологическое, цитологическое, гормональное или ПЦР-исследование; в/м — инъекция внутримышечная; в/в — инъекция внутривенная; в/в кап. — внутривенное капельное введение лекарственных средств.

другими факторами помимо среднего набора медуслуг. Можно отметить, что кластеры среднего набора Middle имеют акцент либо на проведение суточного мониторирования ЭКГ, либо на проведение спирометрии. По времени ожидания выделяется полисомнологическое обследование, которое, вероятно, проводится в последнюю очередь.

Обсуждение

Зашифрованные записи МИС, предоставленные для анализа, содержат как случаи госпитализации по ОМС, так и случаи платной госпитализации, поэтому нельзя говорить о выполнении или невыполнении договорных обязательств по оказанию медицинской помощи по ОМС. Но, как показал анализ, кластеризация данных МИС в среде языка программирования Python может применяться специально для оценки выполнения условий ОМС в медицинских учреждениях с целью выявления удлинения сроков лечения и удорожания стоимости лечения. Стоит отметить, что допускается при наличии показаний проведение дополнительных диагностических исследований, не указанных в стандарте (см. Информационное письмо Минздравсоцразвития России № 14-3/10/2-11668 от 24 ноября 2011 г.). А сопутствующие заболевания лечатся в соответствии со стандартами медицинской помощи с исключением дублирующих услуг.

Пациент с точки зрения социологии медицины является звеном, взаимодействующим с системой здравоохранения и с более широкими социальными системами, выходящими за рамки здравоохранения [5]. Пациент становится частью информатизации общества, концепцию «пациента» можно уже отнести к концепции Intranet-пациента («iПациента»), из чего следует мощная обратная связь с системой здравоохранения через информационные системы. Так, в результате анализа записей МИС выявлен социально-медицинский портрет пациента, страдающего гипертензивной болезнью сердца (ГБ II стадии). Несмотря на то что зашифрованные данные МИС, предоставленные для исследования, не содержали дополнительной информации по сопутствующим заболеваниям, можно сделать важные наблюдения по частоте обследований пациентов.

В табл. 2 представлены нормированные на 100 человек для каждой из двух групп (определенных по распределению всех кластеров) показатели частоты обследований. Группа 1 (стоимость лечения до 35 тыс. руб.) включала 395 (58%) пациентов, средняя арифметическая длительность пребывания в стационаре для этой группы составляло 8,5 сут. Группу 2 (стоимость лечения от 35 тыс. руб.) составили 290 (42%) пациентов, средняя длительность лечения в ней была 11,2 сут.

В группе 1 меньше проведено велоэргометрий и стресс-эхокардиографий, исследований дыхательной системы, сосудов, эндокринной системы, желудочно-кишечного тракта, мочеполовой системы, меньше было эпизодов парентерального лечения. Группа 2 выделялась высокой частотой обследований желудочно-кишечного тракта (эзофагогастродуоденоскопия, колоноскопия, Biohit-тестирование), а также биопсий. В целом среди обследований лидируют ЭКГ, УЗИ бра-

хиоцефальных артерий, брюшной полости, щитовидной железы, эзофагогастродуоденоскопия, комбинированное суточное мониторирование ЭКГ, АД и реопульмонограмма, биопсия и УЗИ малого таза у мужчин.

Можно говорить о социально-медицинском портрете пациента (или iПациента), страдающего гипертензивной болезнью сердца. В основном такому пациенту требуется одно обследование ЭКГ, только в половине случаев проводится внутривенное капельное введение препаратов, почти равное внимание уделяется обследованию сердца и желудочно-кишечного тракта, что оправданно с точки зрения оценки побочных эффектов антигипертензивных препаратов, а также указывает на наличие определенной направленности сопутствующих заболеваний у данной категории больных.

Одним из важнейших показателей, регистрируемых МИС «qMS», является среднее ожидание последнего анализа в минутах. При рассмотрении выборки, ограниченной недельным ожиданием последнего анализа, было выявлено, что тогда срок лечения укладывается в требуемые 5—10 сут. Другой временной показатель — ожидание каждого из обследований — позволил обнаружить, что наименьшее время ожидания пациентом ЭКГ, УЗИ почек и забора крови из пальца связано с кластерами с наименьшим временем пребывания в стационаре.

Выводы

1. Подсчет средних арифметических для оценки работы медицинских учреждений малоинформативен по сравнению с возможностями кластеризации в среде Python. При средней арифметической стоимости лечения 33 262 руб. и длительности госпитализации 9,7 сут кластеризация позволила выявить 2 группы с распределением стоимости лечения до 35 тыс. руб. и свыше 35 тыс. руб. со средней длительностью лечения 8,5 и 11,2 сут соответственно.

2. В условиях информатизации социума возникает понятие «iПациент», характеристики которого отражены в МИС, накапливающих Большие данные от медицинских учреждений. Концепция «iПациент» позволяет анализировать записи МИС с точки зрения социально-медицинской обратной связи и вносить изменения в стандарты лечения в соответствии с обнаруженными реальными требованиями пациентов, формируя гибкие социально-медицинские стандарты «aaS» (as-a-Service, стандарты как услуга).

3. Основываясь на результатах анализа данных МИС, стандарт-aaS в отношении нахождения в стационаре пациентов с АГ может включать обследование не только сердечно-сосудистой системы в сочетании с другими обследованиями с учетом патогенеза заболевания и дифференциальной диагностики гипертензии, но параллельно и обследование ЖКТ с целью раннего выявления сопутствующей патологии и сокращения онкологической смертности населения. По ранее проводимым скрининговым исследованиям в России на поликлиническом уровне было обнаружено, что уже со школьного возраста отмечаются смещения границ популяционной нормы размеров печени по данным УЗИ, что свидетельствует о глубоких социальных причинах

распространения заболеваний ЖКТ в популяции россиян. Высокая частота обследований ЖКТ в данной кардиологической выборке пациентов также указывает на проблему распространенности заболеваний ЖКТ, что требует разработки гибких мер по раннему выявлению заболеваний.

Исследование не имело спонсорской поддержки.

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА

1. Frolio L. Big Data Insights in Healthcare. Part I: Great Ideas Transcend Time. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-healthcare-part-great-ideas-transcend-time>.
2. Frolio L. Big Data Insights in Healthcare. Part II. A Perspective on Challenges to Adoption. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-in-healthcare-challenges-to-adoption>.
3. Frolio L. Big Data Insights in Healthcare. Part III: Mitigating Challenges to Adoption and What Will Follow. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-in-healthcare-part-iii-mitigating-challenges-to-adoption-and-what-will-follow>.
4. Выступление Министра здравоохранения В.И. Скворцовой на расширенном заседании Коллегии Минздрава России. Москва, 20 апреля 2016. Available at: <https://www.rosminzdrav.ru/news/2016/04/20/2903-vystuplenie-ministra-veroniki-skvortsovoy-na-rasshirennom-zasedanii-kollegii-minzdrava-Rossii>.
5. Решетников А.В., Ефименко С.А. *Социология пациента*. М.: Здоровье и общество; 2008.
6. Колесниченко Ю.Ю. *Некоторые аспекты УЗИ в педиатрии на этапе первичной медицинской помощи*. М.: Фолиум; 2008.

Поступила 12.10.2016
Принята в печать 20.01.2017

REFERENCES

1. Frolio L. Big Data Insights in Healthcare. Part I: Great Ideas Transcend Time. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-healthcare-part-great-ideas-transcend-time>.
2. Frolio L. Big Data Insights in Healthcare. Part II. A Perspective on Challenges to Adoption. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-in-healthcare-challenges-to-adoption>.
3. Frolio L. Big Data Insights in Healthcare. Part III: Mitigating Challenges to Adoption and What Will Follow. Dell EMC Global Services Blog «In Focus». Hopkinton MA, USA; 2015. Available at: <https://infocus.emc.com/louis-v-frolio/big-data-insights-in-healthcare-part-iii-mitigating-challenges-to-adoption-and-what-will-follow>.
4. Speech by the Minister of Health of the Russian Federation at the Board Meeting of the Ministry of Health of the Russian Federation. Moscow, 2016. [Vystuplenie Ministra zdravookhraneniya na rasshirennom zasedanii Kollegii Minzdrava Rossii]. Available at: <https://www.rosminzdrav.ru/news/2016/04/20/2903-vystuplenie-ministra-veroniki-skvortsovoy-na-rasshirennom-zasedanii-kollegii-minzdrava-Rossii> (in Russian).
5. Reshetnikov A.V., Efimenko S.A. *The Sociology of the Patient [Sotsiologiya patients]*. Moscow: Zdorov'e i obshchestvo; 2008. (in Russian)
6. Kolesnichenko Yu.Yu. *Some Aspects of Ultrasonography in Pediatrics in Primary Health Care [Nekotorye aspekty UZI v pediatrii na etape pervichnoy meditsinskoj pomoshchi]*. Moscow: Folium; 2008. (in Russian).