



ИПМ им.М.В.Келдыша РАН

Абрау-2016 • Труды конференции



А.А. Захаров, И.Г. Захарова,
Е.А. Оленников, Т.И. Паюсова,
А.В. Бойко

**Интернет-сервис для доказательной
медицины**

Рекомендуемая форма библиографической ссылки

Захаров А.А., Захарова И.Г., Оленников Е.А., Паюсова Т.И., Бойко А.В. Интернет-сервис для доказательной медицины // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2016. — С. 128-134. — doi:[10.20948/abrau-2016-28](https://doi.org/10.20948/abrau-2016-28)

Размещена также [презентация к докладу](#)

Интернет-сервис для доказательной медицины

А.А. Захаров, И.Г. Захарова, Е.А. Оленников, Т.И. Паюсова, А.В. Бойко

ФГБОУ ВО «Тюменский государственный университет»

Аннотация. Приведены определение и основные принципы доказательной медицины. Описаны особенности программно-аппаратной платформы для реализации интернет-сервисов, ориентированных на решение задач доказательной медицины, включая сбор, хранение и интеллектуальный анализ обезличенных медицинских данных. Представлена апробация сервисов в рамках реализации программы сотрудничества университета с Тюменским кардиологическим центром.

Ключевые слова: доказательная медицина, облачный сервис, статистика, нейросетевое моделирование, анализ данных.

1. Введение

Ассоциация организаций по клиническим исследованиям (АОКИ) определяет, что «в основу оценки и признания результатов клинического исследования или скрининга (диспансеризации) должно быть положено высокое качество проводимого исследования, для достижения которого необходимо создать команду профессионалов, где каждый отвечает за свою часть работы» [1]. Участие в таких исследованиях определяется естественной потребностью в профессиональном росте и повышении публикационной активности (в первую очередь магистрантов, клинических ординаторов, аспирантов, докторантов).

Отметим, что такие факторы как принятие de-facto стандартов записи информации, получаемой с лабораторного оборудования DICOM, SCP-ECG, DICOM-ECG и HL7 aECG и др., стандартизация и переход на цифровую форму записи семантической медицинской информации (графические, числовые и текстовые данные, в том числе и записи лечащего врача) [2], наряду с развитием социальных сетей профессионального медицинского сообщества [3-5], способствуют организации совместной работы специалистов разного профиля при проведении медико-биологических исследований.

Тем не менее, в Российском сегменте сети Интернет практически отсутствуют программные продукты, ориентированные на обработку и анализ данных медицинских научных исследований. Среди наиболее популярных в медицинской среде специальных «калькуляторов» — <http://medstatistic.ru> [6], где предлагается платный сервис для сбора данных в таблицы и 15 методов статистической обработки.

2. Проект «Доказательная медицина»

В 1996 г. в работе Дэвида Сэккета и др. [7] впервые было определено понятие «доказательная медицина» (Evidence-Based Medicine, EBM) - как «добросовестное, явное и разумное использование современных научных данных в принятии решений об уходе за отдельными пациентами» (в оригинале «the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients»). В настоящее время с этим понятием связывают методологию организации, проведения и оценки результатов клинических исследований с целью выработки обоснованных рекомендаций по применению конечных результатов исследований [8]. Кроме того, под «доказательной медициной» (ДМ) понимается такая медицинская практика (Evidence-based practice, EBP), когда врач использует при лечении пациента только те методы, результативность которых доказана в исследованиях, соответствующих требованиям рандомизированности и контролируемости [9].

В то же время, с точки зрения информационной основы ДМ, нельзя ограничиваться только использованием результатов рандомизированных контролируемых клинических исследований. Этот принцип находит свое отражение в том, что идеология ДМ основана на учете различных уровней иерархии доказательности тех или иных результатов (EBM pyramid) [10]. В частности, заголовки статей журнала Evidence Based Medicine включают названия этих уровней: Systematic review, Systematic review with meta analysis, Randomised controlled trial, Observational study и др. [11]. Таким образом, развитие ДМ обеспечивается как признанием важности учета личного опыта врачей, так и отказом от абсолютизации результатов, даже имеющих высшую степень доказательности.

Целью Интернет-проекта «Доказательная медицина» является создание масштабируемой информационной системы с аналогичным названием, предоставляющей Интернет-доступ к сервисам сбора, хранения и анализа данных. Под масштабируемостью системы подразумевается масштабируемость по сервисам обработки и по данным: при поступлении данных нового (для системы), но стандартного медицинского формата, имеющиеся сервисы обработки сохраняют свой функционал благодаря работе сервиса преобразования (экстракции) данных.

Пользователями системы «Доказательная медицина» являются врачи-исследователи и специалисты по обработке данных, заинтересованные в качественной обработке результатов клинических исследований и медико-биологических экспериментов.

С точки зрения автоматизации медико-биологических исследований представляется важным тот факт, что *положения ДМ* (наряду с национальным стандартом РФ ГОСТР 52379-2005 «Надлежащая клиническая практика», приказом МЗ РФ №266 от 19.06.2003 г. «Правила клинической практики в РФ»; приказом МЗ РФ №267 от 19.06.2003 г. «Правила лабораторной практики в

РФ», Федеральным законом от 27.07.2006 № 152-ФЗ «О персональных данных», приказом №996 от 05.09.2013 «Об утверждении требований и методов по обезличиванию персональных данных» и другими нормативными документами) *определяют основные стандарты клинических испытаний и принципы надлежащей клинической практики для следующих типов исследований* [12]:

- в одной группе (single group study);
- в параллельных группах (parallel group study);
- в группах "перекрестной" модели (crossover group study).

Отметим, что факт разнородности данных не противоречит существованию стандартов на медицинскую информацию. Это делает возможным не только создание новых, но и использование готовых (для соответствующего семейства типов) сервисов для работы с ними: сервисов хранения и поиска информации, сервисов научной визуализации информации, сервисов интеграции данных, сервисов работы с семантической сетью, сервисов для статистической обработки информации, сервисов интеллектуального анализа информации и др.

Особенности состава и организации взаимодействия указанных сервисов позволяют определить в качестве оптимальной программно-аппаратной платформы для системы «Доказательная медицина» гибридное облако, обеспечивающее при необходимости дальнейшее масштабирование набора соответствующего программного обеспечения в виде сервисов.

3. Гибридное облако ДМ

Преимущества применения облачной технологии определяются не только тем, что с точки зрения пользователей появляется гибкий и удобный инструмент быстрого и удобного доступа к программным приложениям, что повышает эффективность работы исследователей. Дополнительно появляется возможность реализации распределенных проектов для совместного использования программных ресурсов географически разрозненными организациями и группами, что актуально для проектов телемедицины. Наконец, с точки зрения обслуживания системы повышается эффективность контроля за центром обработки данных, снижается время реагирования на события, благодаря инструментам визуализации и анализа контроля над уровнями обслуживания.

Функционирование облака основано на принципе виртуализации программных ресурсов, заключающемся в замене физического представления программных ресурсов логическим. Для организации доступа пользователей к исполняемой среде с программным приложением используются виртуальные машины (VM), которые на физическом уровне представляют собой предварительно подготовленный образ диска, а, по сути, являются контейнером (vApp), в котором развернуто то или иное программное приложение.

В качестве базовой технологии для создания гибридного облака ДМ наилучшие решения с точки зрения простоты масштабируемости предоставляет платформа виртуализации VMware vSphere 5. В качестве модели архитектуры выбрана IaaS, дающая возможность быстрого и удобного изменения масштаба требуемых ресурсов, включая временный доступ к дополнительным виртуальным ИТ-ресурсам для удовлетворения меняющихся требований пользователей [13].

Для непосредственной реализации облачных вычислений была использована технология VMware vCloud Director. Комплексность решений по управлению распределенной программной инфраструктурой в нашем случае обеспечивает предоставление доступа к различным программным приложениям в зависимости от потребностей конкретных пользователей (врачи-исследователи, специалисты по обработке данных).

4. Особенности поддержки деятельности пользователей системы «Доказательная медицина»

В контексте этих требований система ДМ позволяет объединить врачей-исследователей и специалистов по анализу экспериментальных данных из различных медицинских, научных и образовательных организаций. Таким образом, главной особенностью деятельности пользователей при работе с системой является возможность реализации эффективных совместных исследований, опирающихся на единую для всех участников систему сервисов. Закономерно, что такой подход повышает уровень достоверности и качество апробации результатов.

Принципиально важно отметить, что в системе ДМ изначально заложена возможность сбора информации об исследованиях (различного уровня доказательности), проводимых пользователями, в формализованном виде. В качестве основы для структурирования данных нами выбран принцип Evidence Based Cards [14], которые используют модель P.I.C.O. (Patient/Population/Problem, Intervention, Comparison, Outcome). Такой подход позволит в дальнейшем развивать в системе ДМ, например, функционал предметно-ориентированного поиска и агрегации информации по различным признакам (категориям пациентов, диагнозам, методам лечения и др.). Важно заметить, что заполнение соответствующей формы модели P.I.C.O. поможет пользователю формализовать проблему и цель исследования (на уровне представления исходных данных и предполагаемых результатов). Предоставленная информация (размер выборки, число и тип параметров, ретроспективность или проспективность исследования и т.п.) позволит адекватно выбрать необходимые сервисы (параметрические или непараметрические методы для описательной статистики, сравнения выборок, корреляционного анализа и т.д.). На этапе формулировки и уточнения целей исследования осуществляется выбор сервисов для сбора (ввода) экспериментальных данных предоставляющих возможность вести протокол

исследования, фиксировать все действия и необходимую информацию, определять методы, инструменты сбора, форматы и виды требуемых данных. При этом процесс сбора данных основывается на целях и задачах работы и конкретизируется выбранными методами исследования. Предполагается, в случае необходимости, запись повторных замеров с сохранением прошлых результатов в целях предотвращения фальсификаций. Далее определяется структура базы для надежного хранения информации, режимы резервного копирования, архивирования, и определяются технологии предотвращения несанкционированного доступа. Выбираются, а при необходимости разрабатываются, сервисы для анализа данных: визуализация, статистическая обработка, поиск логических закономерностей в собранных данных, построение правил и моделей, объясняющих найденные закономерности и/или прогноз развития процессов с определенной вероятностью.

В качестве возникающих при этом проблем следует выделить случай, когда обезличенные данные необходимо брать из уже имеющихся записей в электронной истории болезни. Это не является сложной задачей, если известна структура базы данных МИС. Если администратор МИС имеет опыт подключения к ЕГИСЗ, то можно организовать автоматическое «попадание» в базу исследования как обезличенных данных новых пациентов, так и выборки по имеющимся данным. В противном случае, можно использовать специальные методы анализа структуры неизвестной базы данных по типу «черного ящика».

Первоначальная регистрация и обсуждение будущего исследования заполнение формы модели P.I.C.O. ведется через сайт проекта. Далее, создается личный кабинет исследователя и с помощью клиентской программы, которая устанавливается на компьютере, пользователь получает защищенный VPN доступ к облаку. Такая технология позволяет создавать необходимые сервисы как win32 – 64 приложения, так и использовать любые уже разработанные сервисы обработки данных для любых операционных систем, которые организованы в виде виртуальных машин (vApp).

5 Апробация проекта

Предлагаемый подход был апробирован при проведении клинического исследования с использованием реальных данных из медицинской информационной системы Тюменского кардиологического центра и Тюменской областной клинической больницы. В проекте участвовало 3 врача-исследователя (в том числе аспирант и докторант) и специалист по обработке данных из университета. Решалась задача определения факторов, влияющих на уровень С-реактивного белка (СРБ), у пациентов со стабильной и нестабильной стенокардией. Подробное описание анализируемых данных представлено в [15].

Были разработаны сервисы, ориентированные на специфические потребности врачей-исследователей: сервис описательной статистики; сервис сравнения выборок, в том числе и для выявления статистически значимых

различий при малом количестве наблюдений; сервис проверки корреляционной зависимости; сервис визуализации и сервисы анализа с помощью нейронных сетей (две архитектуры). Для реализации последних сервисов были построены искусственные нейронные сети с архитектурами MLP (многослойный перцептрон) 81-56-2 (производительность обучения = 98,25) и MLP 74-51-2 (производительность обучения = 100) соответственно для пациентов со стабильной и нестабильной стенокардией. В обоих случаях в качестве алгоритма обучения применялся алгоритм Бройдена-Флетчера-Гольдфарба-Шанно (BFGS). Подробное описание архитектуры нейросетей представлено в [16].

Апробация системы «Доказательная медицина» продемонстрировала эффективность предложенной архитектуры для проведения совместных клинических исследований в части обработки и анализа экспериментальных данных. Однако тема данного пилотного проекта не предполагала использования сложных способов экстракции данных (например, на основе ЭКГ). В настоящее время разрабатывается программное обеспечение соответствующих сервисов на основе алгоритмов машинного обучения с целью интерпретации неформализованной информации.

Литература

1. Ассоциация организаций по клиническим исследованиям (АОКИ). URL: <http://acto-russia.org/> (дата обращения: 10.05.2016).
2. ГОСТ Р 52636-2006 Электронная история болезни. URL: <http://www.resortsoft.ru/publications/gost52636.html> (дата обращения: 10.05.2016).
3. Медицинский видеопортал / Образовательный видеопортал для врачей. URL: <http://med-edu.ru/> (дата обращения: 10.05.2016).
4. «НормаСахар» – онлайн-сервис для пациентов, страдающих сахарным диабетом, и врачей-эндокринологов. URL: <http://normasugar.ru/> (дата обращения: 10.05.2016).
5. Консультант врача. – электронная медицинская библиотека. URL: <http://www.rosmedlib.ru/> (дата обращения: 10.05.2016).
6. Медицинская статистика – сайт для аспирантов и молодых учёных, врачей-специалистов и организаторов, студентов и преподавателей. URL: <http://medstatistic.ru/index.php> (дата обращения: 10.05.2016).
7. Sackett D.L., Rosenberg W.M.C., Muir Gray J.A., Haynes R.B., Richardson W.S. Evidence based medicine: what it is and what it isn't //BMJ. 1996. V. 312. N. 7023. P. 71-72.
8. Власов В.В., Семернин Е.Н., Мирошенко П.В. Доказательная медицина и принципы методологии // Мир медицины. 2001. № 11. С. 22-24.
9. Thyer B.A. What is evidence-based practice? //Brief Treatment and Crisis Intervention. 2004. V. 4. N. 2. P. 167-176.

10. Windish D. Searching for the right evidence: how to answer your clinical questions using the 6S hierarchy // Evid Based Med. 2013. V.18. P.93–97.
11. Evid Based Med. June 2016, Volume 21, Issue 3. URL: <http://ebm.bmj.com/content/current#TherapeuticsPrevention> (дата обращения: 10.07.2016).
12. Основы доказательной медицины. Учебное пособие для системы послевузовского и дополнительного профессионального образования врачей. / Под общей редакцией академика РАМН, профессора Р.Г.Оганова. – М.: Силиция-Полиграф, 2010. – 136 с.
13. Архитектуры облачных систем обработки и хранения данных. URL: <http://www.seagate.com/ru/ru/tech-insights/cloud-compute-and-cloud-storage-architecture-master-ti/> (дата обращения: 10.05.2016);
14. Gray M., Bliss D., Klem M. L. Methods, levels of evidence, strength of recommendations for treatment statements for evidence-based report cards: a new beginning // Journal of Wound Ostomy & Continence Nursing. 2015. V. 42. N. 1. P. 16-18.
15. Захаров А.А., Оленников Е.А., Паюсова Т.И., Петелина Т.И., Мусихина Н.А., Гапон Л.И., Осипова И.В., Такканд А.Г., Белослудцева О.Е. Научный анализ данных в медицинской информационной системе, на примере определения факторов, влияющих на уровень С-реактивного белка, с помощью нейронных сетей // Вестник ТюмГУ. Тюмень: Издательство Тюменского государственного университета, 2014. №7. С. 251-258.
16. Zakharov, A.A., Olennikov, E.A., Payusova, T.I., Silnov, D.S. Cloud service for data analysis in medical information systems using artificial neural networks // International Journal of Applied Engineering Research, 2016. № 11 (4), pp. 2917-2920