

ИСПОЛЬЗОВАНИЕ OLAP ТЕХНОЛОГИИ ДЛЯ ВЫЯВЛЕНИЕ ГРУППЫ РИСКА ОСТРОГО ИНФАРКТА МИОКАРДА

П.А. Зяблецев

Научный руководитель: доцент ОИТ ИШИТР, к.т.н. И.Л. Чудинов
Томский политехнический университет
raz4@tpu.ru

Введение

В настоящее время заболевания сердечно-сосудистой системы являются очень распространенными. Ежегодно от болезней сердца умирают около 17 миллионов человек, что составляет примерно 29% всех летальных случаев [1]. Если посмотреть на статистику в России, то в 2016 году доля смертей от болезней системы кровообращения составляет 47,8% из них 53,3% ИБС (13% смертей от ИБС составляет инфаркт миокарда) [2]. Наиболее распространенным и опасным сердечно-сосудистым заболеванием (ССЗ) является острый инфаркт миокарда (ОИМ), который обычно возникает с развитием острого коронарного симптома (ОКС) так называемого в обиходе сердечного приступа.

Поэтому изучение этой области и вопрос об эффективном выявлении групп риска возникновения ОКС является актуальным.

Проблемы выявления группы риска

Одна из причин высокой смертности от ОИМ – отсутствие эффективных мер по первичной (до проявления симптомов ССЗ) и вторичной (после выявленного ОКС, в том числе ОИМ) профилактике сердечно-сосудистых осложнений, которые обеспечивают своевременное выявление и коррекцию факторов риска (ФР).

Первичная профилактика базируется на статистическом анализе зафиксированных случаях ОКС, в том числе ОИМ вплоть до смертельных случаев. Наиболее распространенным вариантом представления результатов таких исследований является следующая таблица (рисунок 1)

Данная шкала риска «SCORE» разработана экспертами Европейского общества кардиологов на основании данных проспективных исследований, проведенных в 12 странах Европы, в том числе в России (ГНИЦ ПМ), с участием более 205 тысяч больных. Исследования начались с конца 70-х годов и продолжались 27 лет. Рассматривался десятилетний риск развития смертельных случаев всех заболеваний, которые связаны с атеросклерозом. При расчете суммарного риска учитывались два немодифицируемых (пол, возраст) и три модифицируемых ФР (статус курения, систолическое АД, общий ХС).

Данная методика определения риска имеет существенные недостатки:

1) Оценка является приблизительной в связи с тем, что собраны данные из различных регионов из неконтролируемых источников.

2) В данной оценке фигурирует лишь вероятность смерти в 10 летний период, не учитывается вероятность возникновения не фатальных осложнений, а данный фактор очень важен при анализе рисков

3) Состав факторов риска и степень их влияния на возникновение серьезных сердечно - сосудистых случаев (СССС) может меняться со временем и регионом проживания человека. Диаграмма учитывает лишь 5 факторов. Изменить состав факторов невозможно. В то время как со временем отношение к наиболее важным факторам меняется (например, индекс массы тела, сопутствующие заболевания, регион проживания, ранее перенесенные СССС и др.)

4) «Визуально - ручной» способ определения оценки риска СССС при котором затруднен оперативный анализ оценки риска при целевых изменения значений факторов (снижение веса, прекращение курения и т.п.)

Созданию информационной системы с аналогичными функциями, но лишенной перечисленных недостатков и посвящена настоящая работа. Основная особенность – ориентация на региональный аспект исходных данных и в частности на данные Томского регистра острого инфаркта миокарда, который ведется с 1984 года (всего учтено около 50 000 случаев ОКС, в том числе более 25 000 подтвержденных случаев острого инфаркта миокарда).

В ряде исследований показано, что прогноз развития и течения сердечно-сосудистых заболеваний значительно хуже при сочетании нескольких, даже умеренно выраженных ФР по сравнению с одним высоким ФР.

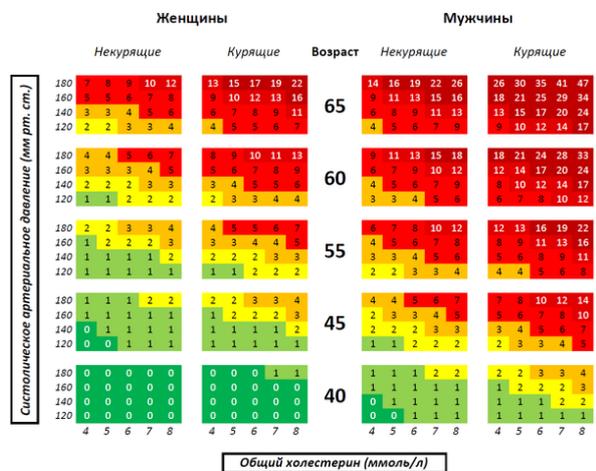


Рис. 1. Шкала SCORE

Если человек сможет оценить опасность возникновения ОКС (в том числе ОИМ) и предпринять необходимые меры для снижения этой вероятности, то это может спасти множество жизней. По статистике около половины пациентов, перенесших ОИМ, умирают в течении последующих трех лет. Часть людей после инфаркта или инсульта не могут вернуться к полноценной жизни из-за их серьезных последствий. Поэтому эффективность профилактических мер по предотвращению ССЗ невозможно переоценить.

Пути решения проблемы

Для анализа всех факторов, которые могут влиять на риск возникновения ОКС и ОИМ целесообразно использовать OLAP технологию. Определенные группы риска традиционно базируется на выполнении запросов к базе данных случаев ОКС. Типовая структура запроса – это определение числа случаев ОКС определенного типа (в том числе смертельных от ОИМ) для фиксированных значений параметров, влияние которых на возникновение случая ОКС необходимо определить. Например, систолическое давление, возраст, пол, наличие пристрастия к курению, уровень холестерина, наличие атеросклероза и т.д. Более детальный чем в SCORE анализ (более детальные интервалы и увеличение числа параметров) требует выполнения нереально большого числа запросов равного сочетанию чисел возможных значений каждого из учитываемых параметров.

Если для повышения точности интервалы делать более детальными и учитывать другие не менее важные параметры, такие как индекс массы тела, регион, сахарный диабет, атмосферное давление и другие, то количество запросов будет увеличиваться в геометрической прогрессии. Таким образом число запросов при организации хранения и реализации запросов традиционным способом (РМД и SQL) становятся не реальными. Именно это обуславливает необходимость применения OLAP технологии для определения групп риска ОКС. Основная идея OLAP заключается в предварительной обработке статистических данных (в нашем случае о фактах ОКС) путем создания так называемого «куба»

Пример «куба» OLAP для трех параметров (давление, возраст, уровень холестерина) приведен на рисунке 2. В каждой ячейке храниться количество смертей от ОИМ на 1000 человек (мера куба) для конкретных значений (интервалов значений) давления, возраста и уровня холестерина. Наглядно изобразить «кубы» для более чем трех параметров не представляется возможным.

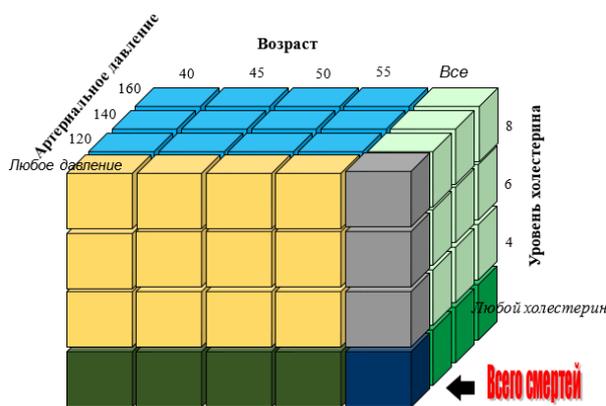


Рис. 2. Пример OLAP-куба нашей предметной области

Для построения «куба» вначале создается специальным образом организованное хранилище данных по выбранному множеству параметров (атрибутов) исходных статистических данных, а «кубы по любой комбинации параметров формируются из хранилища эффективным способом. Из «куба», используя инструмент OLAP, можно оперативно получить таблицы, иллюстрирующие зависимости между значениями любых параметров.

Заключение

Определение группы риска ОИМ можно определить на основе статистических данных о реальных случаях ОКС, например, на основе Томского регистра ОИМ. Применение традиционного подхода (РМД и SQL) к определению статистических зависимостей между параметрами, зафиксированными в регистре случаев ОКС практически невозможно. Использование технологии OLAP позволяет значительно повысить эффективность статистической обработки для выявления групп риска ОИМ за счет специальной организации хранения исходных статистических данных (хранилища и «кубы») и инструментария OLAP для «куба»

Список использованных источников

1. Сердечно-сосудистые заболевания – основная причина смертности на Земле [Электронный ресурс]. – URL: <http://www.kardi.ru/ru/index/Article?Id=14&ViewType=view> (Дата обращения: 22.10.2018г.)
2. Здравоохранение в России. 2017: Стат.сб./Росстат. - М., 2017. – 170 с. ISBN 978-5-89476-448
3. OLAP Council. [Электронный ресурс]. – URL: www.olapcouncil.org (Дата обращения 01.11.18г.)