Подходы к оценке качества наблюдательных исследований клинической практики, основанных на анализе «больших данных»

Гиляревский С.Р.*

Российская медицинская академия непрерывного профессионального образования, Москва, Россия

Статья посвящена обсуждению проблем оценки качества обсервационных исследований реальной клинической практики и определению их места в иерарахии доказательной информации. Рассматривается понятие «большие данные» и приемлемость использования такого термина для обозначения крупных обсервационных исследований. Приводятся данные об ограничениях административных баз данных и баз данных страховых заявок при выполнении обсервационных исследований для оценки эффектов вмешательств. Обсуждается понятие вмешивающихся факторов, влияющих на результаты обсервационных исследований. Рассматриваются современные подходы к уменьшению выраженности систематически ошибок в исследованиях реальной клинической практики. Приводятся критерии оценки качества обсервационных фармакоэпидемиологических исследований и принципиальные отличия от рандомизированных клинических исследований. Обсуждаются результаты систематических обзоров исследований реальной клинической практики для оценки эффектов применения прямых пероральных антикоагулянтов.

Ключевые слова: исследования реальной клинической практики, «большие данные», фармакоэпидемиологические исследования, прямые пероральные антикоагулянты.

Для цитирования: Гиляревский С.Р. Подходы к оценке качества наблюдательных исследований клинической практики, основанных на анализе «больших данных». *Рациональная Фармакотерапия в Кардиологии* 2021;17(4):584-593. DOI:10.20996/1819-6446-2021-08-01.

Approaches to Assessing the Quality of Observational Studies of Clinical Practice Based on the Big Data Analysis Gilyarevsky S.R.*

Russian Medical Academy of Continuing Professional Education, Moscow, Russia

The article is devoted to the discussion of the problems of assessing the quality of observational studies in real clinical practice and determining their place in the hierarchy of evidence-based information. The concept of "big data" and the acceptability of using such a term to refer to large observational studies is being discussed. Data on the limitations of administrative and claims databases when performing observational studies to assess the effects of interventions are presented. The concept of confounding factors influencing the results of observational studies is discussed. Modern approaches to reducing the severity of bias in real-life clinical practice studies are presented. The criteria for assessing the quality of observational pharmacoepidemiological studies and the fundamental differences between such studies and randomized clinical trials are presented. The results of systematic reviews of real-life clinical trials to assess the effects of direct oral anticoagulants are discussed.

Key words: clinical practice studies, big data, pharmacoepidemiological studies, direct oral anticoagulants.

For citation: Gilyarevsky S.R. Approaches to Assessing the Quality of Observational Studies of Clinical Practice Based on the Big Data Analysis. *Rational Pharmacotherapy in Cardiology* 2021;17(4):584-593. DOI:10.20996/1819-6446-2021-08-01.

*Corresponding Author (Автор, ответственный за переписку): sgilarevsky@rambler.ru

«Большие данные» и «данные клинической практики» как возможный источник доказательной информации

Термин «данные реальной клинической практики» (ДРКП) нередко используют для обозначения обсервационных, или наблюдательных, исследований, основанных на анализе так называемых «больших данных» — «big data» (БД). Отсутствует согласованное мнение по поводу определения ДРКП, но обычно к ним относят «данные о состоянии здоровья пациентов и/или предоставлении медицинской помощи, которые получают из разных источников» [1]. На использование

пусть и условно относят к БД, обычно служат электронные медицинские карты (ЭМК), административные базы данных или базы данных страховых компаний [3].

БД возлагают большие надежды как на источник по-

лучения научно обоснованных ДРКП. Впрочем, следует

отметить, что формально медицинские данные, которые

анализируют как ДРКП, в большинстве случаев нельзя

относить к БД. Пороговый объем информации, который

определяют как БД, измеряется в терабайтах (10^{12} байт), петабайтах (10^{15} байт) или даже экзабайтах

Источниками медицинской информации, которые

Основной характеристикой БД считается объем данных и их сложность. Обычно при описании сложности БД используют три английские буквы «V»: volume,

Received/Поступила: 16.08.2021 Accepted/Принята в печать: 20.08.2021 (1018 байт) [2].

variety, velocity, что, помимо большого объема данных, означает их разнообразие и скорость обновления [2,4,5]. Термин БД используют для описания больших массивов информации, которые делают трудными обработку баз данных вследствие сочетания их большого размера, высокой частоты обновления или ошибок, обусловленных действием дополнительной неоднородности [4]. Следует отметить, что БД и ДРКП нельзя считать синонимами, однако ДРКП могут использовать БД, и важно определить условия, при которых их использование при анализе ДРКП позволяет сделать результаты источником доказательной информации.

Однако очевидно, что рандомизированные клинические исследования (РКИ) остаются стандартным методом получения научных доказательств эффективности вмешательств, и в настоящее время неясно, могут ли результаты обсервационных исследований хотя бы отчасти отвечать на те же вопросы, на которые дают ответ результаты крупных и хорошо организованных РКИ [3].

Использование рандомизации при оценке эффективности вмешательства считается одним из основных принципов доказательной медицины [6]. Нерандомизированные, т.е. обсервационные (или наблюдательные) исследования, в первую очередь, с проводят с описательными целями. Они имеют ограниченные возможности для сравнения эффектов вмешательств в связи с невозможностью надежно устранить вмешивающиеся факторы («confounders»), то есть, возможные источники систематических неизмеренных (или в принципе неизмеримых) факторов, если такие вмешивающиеся факторы имеют связь с изучаемым воздействующим фактором при отсутствии причинно-следственной связи с ним, но при наличии причинно-следственной связи с изучаемым клиническим исходом [7]. О влиянии вмешивающегося фактора на результаты исследования говорят в тех случаях, когда кажущаяся очевидной причинно-следственная связь между изучаемым вмешательством и клиническим исходом, по крайней мере, отчасти, обусловлена влиянием третьего (вмешивающегося) фактора [8]. С целью уменьшения различий между группами в обсервационных исследованиях по учтенным характеристикам используют несколько методов анализа, но ни один из них не позволяет учитывать неизмеренные, или в принципе неизменяемые характеристики участников исследования [6]. Следует отметить, что в целом примерно для 90% клинических рекомендаций отсутствуют подтверждения, полученные в ходе выполнения крупных РКИ с высоким качеством [9], поэтому понимание ограниченных доказательных возможностей обсервационных исследований представляется важным.

Вопросы к статистическим методам и качеству данных, полученных в исследованиях клинической практики

Обычно при оценке приемлемости использования БД в обсервационных исследованиях необходимо ответить на три вопроса: 1) соответствуют ли данные целям исследования; 2) соответствует ли качество данных целям исследования; 3) приемлемо ли выполнение более подробного статистического анализа, особенно в тех случаях, когда предполагаемая точность выводов завышена [3].

Во-первых, большие базы данных могут недостаточно отражать популяцию, которая представляет интерес, а большое число данных не означает, что встречаемость определенных факторов или абсолютный риск могут быть оценены без систематической ошибки. Отсутствие репрезентативности нередко вызывает меньшую озабоченность, когда основное внимание уделяется установлению относительных связей, т.е. отношению рисков. Во-вторых, наличие большого массива данных об индивидуальных характеристиках пациентов не устраняет необходимость измерения систематических ошибок и оценки пропущенных данных

Качество данных, собираемых в условиях клинической практики, должно особенно тщательно проверяться [3]. Это обусловлено, например, тем, что ЭМК заполняют разные врачи, и такие записи могут различаться по подробности и точности. Более того, по разным причинам некоторые данные могут быть пропущены, а сами данные отражают лишь те показатели, которые регистрируют в клинической практике [10]. Не так давно стали учитывать и данные, предоставляемые самими пациентами, включая оценку ими симптомов заболевания в условиях повседневной жизни. Надежность таких данных подвергают сомнению в связи с возможностью систематических ошибок. Например, такие ошибки могут быть обусловлены социальными предпочтениями пациента (например, указанная им частота приема препарата может быть некорректной), а также неточностями воспоминания, ошибками измерения определенных показателей, например, измерения АД в домашних условиях [11,12].

Несмотря на то, что оценка большого числа показателей, получаемых в клинической практике, позволяет учитывать множество вмешивающихся факторов, анализ таких факторов представляет собой самостоятельную и очень непростую проблему [13]. К сожалению, при проведении популяционных эпидемиологических исследований, предпринимаемых для оценки большого числа действующих на пациента факторов и оценки частоты развития клинических исходов, может не регистрироваться подробная информация, отражающая

все характеристики пациента, и может отсутствовать детальная информация о вмешательствах, клинических исходах или вмешивающихся факторах [14].

Считается, что получение доказательной информации, основанной на анализе ДРКП, должно регулироваться тщательно разработанными рекомендациями, а выполнение анализа — тщательно контролироваться для обеспечения точных и неискаженных систематическими ошибками результатов [15].

Необходимость рекомендаций по проведению фармакоэпидемиологических исследований

Одна из главных трудностей интерпретации фармакоэпидемиологических исследований состоит в невозможности точного учета вмешивающегося фактора как показания к применению препарата. Другая проблема состоит в точности данных, собираемых в условиях клинической практики. Особенно это касается информации, полученной в базах данных страховых компаний или баз данных, созданных для выставления счетов за медицинские услуги (например, в так называемых административных базах данных). Ограниченность такого источника информации, используемого для выполнения фармакоэпидемиологического исследования, заключается в трудности идентификации участников исследования и в оценке истинной экспозиции препарата, а также в наличии вмешивающихся факторов и возможности учета неблагоприятных клинических исходов только на основании кодов. В таких случаях следующая важная информация может быть либо ограничена, либо недоступна: показания к применению препарата, личные предпочтения пациента, роль врача и пациента в выборе определенной тактики лечения, а также наличие любых противопоказаний к использованию определенного лекарственного средства и тяжесть заболевания. Все указанные вмешивающиеся факторы могут искажать полученные результаты оценки связи между применением определенного препарата и частотой развития изучаемого неблагоприятного клинического исхода. Даже в случае использования клинических данных (например, записей в ЭМК) показания или противопоказания к назначению препарата могут не регистрироваться, или такие сведения могут содержаться в области для свободного заполнения текстовой информации (и не подлежат формализованному автоматическому анализу), что исключает их использование при выполнении фармакоэпидемиологических исследований. Таким образом, информация об очень важных вмешивающихся факторах может быть недоступна как для исследователей, так и для врачей. Поэтому авторы исследований должны сообщать о возможных вмешивающихся факторах при использовании данных,

полученных в условиях клинической практики, а также следует отмечать, какие подходы были предприняты к уменьшению влияния таких факторов на полученные результаты [16].

Рекомендуется включать в раздел статьи «обсуждение» или в «выводы» текст с уточнением того, могут ли полученные результаты быть объяснены вмешивающимися факторами, связанными с предпочтительным назначением вмешательства. Такие уточнения помогут уменьшить возможность неправильного использования представленной в статье информации для принятия клинического решения и повысить доверие к полученной доказательной информации и ее интерпретации. Такие пояснения могут содержать данные о любом вторичном анализе, выполненном для оценки надежности полученных результатов и альтернативном их объяснении. Например, для уточнения того, различались ли пациенты изучаемых групп по предпочтительности назначения определенного вмешательства при наличии сходных заболеваний. Если анализировали базы данных страховых компаний, то важно указать, учитывалась ли дата регистрации пациента в базе данных, и обсудить влияние на результаты исследования пропущенных данных о ранее применявшемся лечении или ранее развившихся исходах.

Формализованная оценка качества фармакоэпидемиологических исследований

Разработано довольно много рекомендацией по проведению обсервационных исследований, но отсутствуют специальные рекомендации по проведению исследований, основанных на анализе ДРКП, результаты которых используют регуляторные органы для одобрения определенных показаний к применению лекарственных препаратов. Широко используемые рекомендации по подготовке публикаций с отчетом о результатах обсервационных исследований включают такие, как RECORD (REporting of studies Conducted using Observational Routinely collected health Data) [17] и STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) [18]. Рекомендации RECORD были основаны на систематическом обзоре наиболее существенных недостатков в отчетах о результатах обсервационных исследований, в которых использовали информацию, полученную в условиях клинической практики [19].

В современных рекомендациях по правилам сообщений результатов нерандомизированных исследований обычно содержится контрольный список (checklist), который позволяет избежать наиболее частых ошибок в отчетах о подобных исследованиях. Очевидно, что с помощью такого списка можно оце-

нивать и качество уже опубликованных исследований. В частности, в рекомендациях RECORD содержится контрольный список, состоящий из 13 вопросов. После опубликования документа он был одобрен более чем 20 центральными медицинскими журналами.

Методологическая сложность фармакоэпидемиологических исследований определяет потребность в более строгих и специфичных рекомендациях для выполнения фармакоэпидемиологических исследований по сравнению с рекомендациями RECORD и STROBE. С этой целью были разработаны специфичные рекомендации RECORD-PE (REporting of studies Conducted using Observational Routinely collected health Data Pharmacoepidemiological Research) [16] для оценки качества фармакоэпидемиологических исследований. Следует отметить, что, в отличие от ранее предложенных соответствующих рекомендаций, в которых основное внимание уделялось оценке качества отчетов о результатах обсервационных исследований, в RECORD-PE акценты делаются на принятых методологических подходах к выполнению фармакоэпидемиологических исследований и оценке качества публикаций с отчетом об их результатах [20-22].

Контрольный список для оценки качества фармакоэпидемиологических исследований в рекомендациях RECORD-PE

Контрольный список для оценки качества фармакоэпидемиологических исследований, разработанный в ходе подготовки рекомендаций RECORD-PE, создавался при участии международных экспертов, которые в ходе регулярных телеконференций и с помощью электронных сообщений вносили поправки в исходный документ. Окончательный вариант был принят после очного обсуждения на совещании 25 августа 2017 г. Для включения в окончательный контрольный список каждого предлагаемого пункта необходимо было согласие более 80% экспертов по поводу его содержания и формы. Кроме того, до принятия окончательного варианта контрольного списка в него были внесены поправки, сделанные членами Международного общества специалистов по фармакоэпидемиологии.

В целом контрольный список рекомендаций RECORD-PE включает 22 пункта. Остановимся подробнее на тех пунктах контрольного списка, которые отличаются от ранее предложенных в рекомендациях RECORD и STROBE (нумерация как в оригинале). В целом таких пунктов было 7 (табл. 1).

Таким образом, знакомство с указанными в блоке критериями оценки позволяет предположить трудность соответствия всем указанным в рекомендациях RECORD-PE требованиям к фармакоэпидемиологическому исследованию.

Административные базы данных и базы данных страховых заявок как источник данных в фармакоэпидемиологических исследованиях

Формируемые проспективно базы данных, например, базы данных страховых заявок или записей в ЭМК наиболее часто используют в качестве источника информации для проведения исследований, цель которых состоит в оценке особенности использования препарата в условиях клинической практики, а также для получения данных об эффективности и безопасности применения его в таких условиях. Транзакционные базы данных предоставляют проспективно накопленную информацию о применяемых методах лечения и клинических исходах у миллионов пациентов. Считается, что размер таких баз данных продолжает неуклонно увеличиваться, содержащаяся в них информация становится более детализированной и доступной за счет установления связей между данными, их стандартизации и возможностью широкого использования [23]. Очевидно, что использование таких проспективно сформированных баз данных позволяет оценивать эффекты применения лекарственных средств в группах пациентов, которые не были включены в РКИ, например у беременных, пациентов очень пожилого возраста или имеющих большое число сопутствующих заболеваний, что позволяет точнее отражать клиническую практику использования препарата и его действие в условиях полифармакотерапии. Наконец, такие исследования могут быть выполнены в более короткие сроки и при меньших финансовых затратах по сравнению с РКИ.

Регуляторные органы в течение многих десятилетий используют такие базы данных для оценки безопасности применения медицинских продуктов, и в некоторых случаях результаты таких исследований учитываются при принятии решения об одобрения препаратов [24]. Несмотря на это, надежность анализа ДРКП подвергается сомнениям [25]. Эксперты, принимающие решение о тактике применения препаратов, отдают предпочтение РКИ не только в связи с исходной рандомизацией, но также потому, что в ходе выполнения РКИ имеется возможность тщательного наблюдения за точностью измеряемых показателей, определенных характеристик пациентов и развившимися клиническими исходами. Правильный протокол РКИ и его тщательное выполнение обеспечивает возможность установить причинно-следственные связи между вмешательством и риском развития клинического исхода, что принципиально важно для принятия решения об одобрении вмешательства регуляторными органами. Напротив, анализ ДРКП в отсутствие рандомизации и намеренного сбора данных часто воспринимается как сложный, недостаточно понятный и плохо представленный источник

Table 1. Pharmacoepidemiological study specific items on the RECORD-PE checklist that have been added to the earlier RECORD guidelines [16]

Таблица 1. Специфичные для фармакоэпидемиологических исследований пункты контрольного списка RECORD-PE, которые были добавлены в принятые ранее рекомендации RECORD [16]

Nº	Содержание пункта			
4.a	Включено ли детальное описание схемы исследования (и его характеристик), а также сообщалось ли о наличии нескольких целей исследования, если таковые были.			
4.b	Рекомендуется использование диаграмм для отражения ключевых аспектов схемы исследования, включая экспозицию препарата, период вымывания, период задержки в начале применения препарата и период наблюдения, а также при определения ковариат, если приемлемо.			
6.1.a	Описаны ли критерии включения в исследование и порядок применения таких критериев для установления популяции, в которой предполагается выполнение исследования. Есть ли указание на то, что в исследование включали пациентов, имеющих определенные показания к назначению препарата и допускалось ли только однократное или и многократное включение пациентов в исследование.			
7.1.a	Опишите, как разрабатывалось определение экспозиции препарата.			
7.1.b	Укажите источники информации, с помощью которых получали данные об влиянии препарата у пациентов.			
7.1.c	Опишите период, в течение которого считалось, что пациент применяет препарат или препараты. Необходимо предоставить обоснование по выбору определенного периода, в течение которого пациент применяет препарат. Необходимо определить критерии усечения и цензурирования данных слева (для пациентов, которых можно было включить в исследование после определенной даты, но которые не были включены в исследование).			
7.1.d	Установите, как развившиеся неблагоприятные исходы связаны с влиянием препарата в настоящее время, в прошлом или с кумулятивным действием препарата.			
7.1.e	При оценке связи между дозой препарата и риском развития неблагоприятного исхода следует описать, каким образом учитывали применение препарата в настоящее время или ранее, а также продолжительность периода, в течение которого пациент принимал препарат.			
7.1.f	Использование в исследовании любой группы сравнения должно быть отмечено и обосновано.			
7.1.g	Следует отметить, какой подход будет использован для учета пациентов, у которых в ходе выполнения исследования применяли более одного исследуемого препарата.			
8.a	Опишите систему здравоохранения и механизм формирования записей о влиянии препарата. Укажите тип медицинского учреждения, в котором назначали исследуемый препарат.			
12.1.a	Опишите методы, которые использовали для оценки сделанных допущений. (Это необходимо сделать, так в фармакоэпидемиологических исследованиях каждый пациент становится контролем самому себе. В таких случаях необходим ряд допущений для обеспечения надежности данных и отсутствие систематических ошибок, например для того, чтобы исключить влияние ранее перенесенного неблагоприятного клинического исхода).			
12.1.b	Опишите и обоснуйте использование схемы исследования для оценки эффектов нескольких препаратов, а также определенных характеристик схемы исследования и подходов к анализу данных.			
19.1.a	Опишите, в какой степени выбранная база данных соответствует цели получения информации о влиянии исследуемого препарата.			
20.a	Обсудите возможные вмешивающиеся факторы, обусловленные систематическими ошибками, связанными с показаниями и противопоказаниями к назначению препаратов или тяжестью заболевания, а также отбором пациентов (например, более здоровые могут быть более приверженными к терапии, а у более тяжелых имеется более высокая вероятность прекращения терапии) с целью альтернативного объяснения результатов исследования, если приемлемо.			

информации, который трудно интерпретировать и воспроизвести. Несмотря на расширяющиеся возможности использования ДРКП в качестве дополнительного источника информации об эффектах применения препаратов, а также частую подмену таким источником результатов РКИ с целью продвижения препаратов на фармацевтическом рынке, отсутствие надежности данных, полученных в отсутствие рандомизации при анализе ДРКП, ограничивает ценность такого источника доказательной информации.

Ненадежность базы данных страховых заявок для измерения риска развития кровотечений у пациентов, применяющих антитромботические препараты после развития острого коронарного синдрома, была убедительно продемонстрирована с помощью анализа частоты развития неблагоприятных клинических ис-

ходов у участников проспективного обсервационного исследования TRANSLATE-ACS, информация о которых также была включена в базу данных страховых заявок [26]. Результаты сопоставления частоты развития кровотечения, которую оценивали в ходе проспективного исследования и с помощью информации, содержащейся в базе данных страховых заявок, свидетельствовали о существенных различиях по частоте развития такого исхода. В целом в анализ были включены данные о 12365 пациентах (71,9% мужчины; средний возраст $60\pm11,6$ года), которые перенесли острый коронарный синдром и принимали антиагреганты. Кумулятивная частота развития кровотечений в течение 1 года, которая основывалась на данных, содержащихся в страховых заявках, и по данным, полученным в исследовании TRANSLATE-ACS, составляла 5,0% и 5,4%

соответственно. Согласованность результатов анализа, выполненного с помощью базы данных страховых компаний и базы данных проспективного исследования TRANSLATE-ACS для частоты развития умеренно выраженных или тяжелых кровотечений по шкале GUSTO, была очевидно плохой с коэффициентом k для k статистики 0,24 (95% доверительный интервал [ДИ] от 0,11 до 0,20).

Сравнение результатов анализа данных клинической практики и рандомизированных клинических исследований

При сравнении результатов анализа ДРКП с результатами РКИ многие авторы делают вывод о схожести полученных в таких исследованиях результатах [27-29]. В то же время в отдельных случаях могут быть существенные различия клинической практики. Несмотря на то, что РКИ и исследования клинической практики часто несколько различаются по вопросам, на которые предполагают получить ответ, такое несовпадение результатов вызывает опасения экспертов [23]. В некоторых случаях результаты различаются не только по выраженности установленного эффекта, но и по направлению, что приводит к противоположным выводам об эффектах препарата [30,31]. Такие случаи могут быть обусловлены ошибками в протоколе исследований или выполненном анализе, так что они в целом могут быть предотвращены при соблюдении методологических подходов. Остановимся на наиболее часто упоминаемом случае расхождения выводов обсервационного исследования и РКИ.

Начальные результаты обсервационного исследования NHS (Nurses' Health Study) указывали на наличие статистически значимой связи между применением постменопаузальной гормональной терапии (ПМГТ) и менее высоким риском развития осложнений ССЗ [32], но противоречили полученным позднее в ходе выполнения РКИ WHI (Women's Health Initiative) данным [31]. Однако результаты повторного анализа данных об участниках исследования NHS позволили устранить выявленные противоречия и установить наиболее вероятную причину несовпадения результатов обсервационного и рандомизированного исследования [33]. Оказалось, что при выполнении исходного анализа данных об участниках исследования NHS с помощью стандартного регрессионного анализа исключались данные о женщинах с коронарной болезнью сердца, которая развилась в ранние сроки после начала ПМГТ, т. е., была допущена корректируемая систематическая ошибка, которую в фармакоэпидемиологии принято обозначать как «истощение предрасположенности», или в других исследованиях как «систематическую ошибку, связанную с выживаемостью» [34]. Учет исходных характеристик участников исследования, которые применяли или не применяли ПМГТ, не позволяет преодолеть такую систематическую ошибку, но уточнение клинического вопроса и правильное планирование и выполнение повторного анализа позволяет это сделать. В любом случае такой пример подчеркивает необходимость соблюдать осторожность при трактовке результатов обсервационных исследований и критически относится к их выводам.

В то же время есть примеры, когда тщательный учет возможных систематических ошибок и правильный план анализа позволил получить в исследованиях клинической практики результаты, сходные с таковыми в РКИ. Например, в ходе выполнения исследования, основанного на анализе базы данных для выставления счетов в стационаре на оплату коронарного шунтирования, было отмечено увеличение риска смерти при применении апротинина по сравнению с использованием аминокапроновой кислоты [35]. Опубликованные в том же году, но несколько позднее, результаты РКИ BART (Blood Conservation Using Antifibrinolytics in a Randomized Trial) подтвердили увеличение такого риска, что обусловило удаление апронитина с фармацевтического рынка США [36]. В табл. 2 представлены основные характеристики РКИ и обсервационных исследований клинической практики.

Подходы к уменьшению выраженности систематических ошибок в исследованиях клинической практики

Обсервационные исследования почти всегда характеризуются наличием систематических ошибок, поскольку в таких исследованиях прогностические факторы неравномерно распределены между пациентами, у которых применяется разная тактика лечения. Стандартным подходом к решению такой проблемы становится выполнение анализа с учетом потенциально вмешивающихся факторов или стратифицированного анализа (табл. 3). Такой подход применяют для измерения выраженности факторов риска с целью создания прогностически однородных групп и объединения установленного эффекта для групп пациентов, имеющих различные характеристики [37].

Альтернативой такому стандартному подходу может быть анализ, выполненный с учетом предпочтительного применения определенного вмешательства (propensity analysis). В ходе выполнения такого анализа группы пациентов, у которых применяется определенный метод лечения, формируются таким образом, чтобы они были сопоставимы по степени предпочтительного назначения определенного вмешательства (см. табл. 3). Применение такого метода позволяет учитывать действие многих факторов риска даже при развитии неблагоприятных исходов у относительно небольшого

Table 2. Comparison of randomized controlled trials and observational studies of clinical practice Таблица 2. Сравнение рандомизированных контролируемых исследований и обсервационных исследований клинической практики

Характеристика	Рандомизированные контролируемые исследования	Исследования «реальной клинической практики»	
Тип исследования	Экспериментальное/интервенционное	Обсервационное/неинтервенционное	
		Ретроспективное/проспективное	
Основные задачи	Оценка эффективности, безопасности, экономической эффективности	Оценка эффективности, безопасности, экономической эффективности, естественного течения, согласия пациентов соблюдать предписанный режим терапии и приверженность к нему, моделей оказания медицинских услуг, предпочтения пациентов.	
Популяция пациентов	Ограничена пациентами, имеющими строго определенные характеристики и мотивацию к участию в исследовании.	Большая популяция пациентов, имеющих различные характеристики, не ограниченные строгими критериями.	
Наблюдение	Интенсивное наблюдение в соответствии с принципами ICH-GCP).	В большинстве случаев не требуется	
Препарат/вмешательство	Плацебо (золотой стандарт) или активный	Отсутствует или стандартная клиническая практика, может оцениваться	
сравнения	препарат сравнения.	большое число препаратов/вмешательств.	
Оцениваемые исходы	Четкая последовательность оцениваемых исходов	Широкий диапазон оцениваемых исходов.	
Вмешивающие факторы, связанные с получением информации	Стандартизованный и контролированный сбор данных.	Использование обычных данных, которые регистрируют в клинической практике; наличие систематических ошибок, связанных с отбором, неточными воспоминаниями пациентов при их интервьюировании.	
Рандомизация	Да	Нет	
Возможность использования слепого метода	Да	Нет	
Наблюдение	Обычно относительно непродолжительное.	Отражает применяемую в клинической практике.	
ICH-GCP – International Conference on H	larmonisation of Good Clinical Practice		

числа пациентов. В то же время следует отметить, что propensity analysis не лишен других недостатков стандартных подходов к учету вмешивающихся факторов, так как и при выполнении такого анализа исследователи могут не измерить (или измерить неточно), или не учесть все значимые прогностические факторы, поэтому влияние таких неучтенных или неизвестных вмешивающихся факторов может привести к систематическим ошибкам и искажению результатов исследования [37].

Еще одним способом для уменьшения систематических ошибок в обсервационных исследованиях считается использование метода анализа инструментальных переменных, основу которого составляет анализ с учетом переменных, связанных с вероятностью применения определенного вмешательства, но не связанных с любым прогностическим фактором или клиническим исходом (за исключением исхода, на который влияет такое вмешательство). Считается, что указанный подход по эффекту может напоминать рандомизацию, и он, как и другие рассмотренные методы учета вмешивающихся факторов, не устраняет возможность систематических ошибок (см. табл. 3) [37].

Таким образом, все указанные методы позволяют уменьшить вероятность систематических ошибок обсервационных исследований, ни один из таких методов не позволяет добиться сбалансированности групп не только по известным, но и по неизвестным или неустановленным прогностическим факторам, т. е., добиться такой сбалансированности, которая обеспечивается рандомизацией.

Наконец, обсуждая проблемы улучшения качества выполнения обсервационных исследований, результаты которых могут влиять на принятие решений в здравоохранении, и подходы к уменьшению систематических ошибок нельзя не упомянуть международную профессиональную некоммерческую организацию ISPOR (International Society for Pharmacoeconomics and Outcomes Research) [38]. Задача этой организации состоит в улучшении качества фармакоэкономических исследований и исследований по оценке влияния медицинских вмешательств на благополучие пациента, которое оценивается по клиническим и экономическим показателям, а также по критериям, отражающим потребности пациента. Главной задачей организации стало усовершенствование принятие решений в здравоохранении за счет улучшения качества клинических исследований. К началу 2021 г. число членов ISPOR достигало 15000 человек из более чем 100 стран

Table 3. Main methodological approaches to reducing the imbalance of prognostic factors between groups Таблица 3. Основные методологические подходы к уменьшению дисбаланса прогностических факторов между группами

Тип анализа	Описание метода	Основные преимущества	Основные недостатки
Традиционный регрессионный анализ с учетом факторов риска	С помощью анализа оценивают связь между выбором метода лечения и изучаемым клиническим исходом. Выравнивание групп по прогностическим факторам путем создания прогностически однородных страт и комбинирования результатов по слоям и объединение результатов по стратам. С помощью многофакторного анализа, выполненного с учетом нескольких прогностических факторов, рассчитывается один стандартизованный показатель эффективности лечения, в котором учтены указанные факторы.	Давно принятый и широко применяемый метод учета вмешивающихся факторов в обсервационных исследованиях. Интерпретация положительных и отрицательных эффектов лечения проще, чем при использовании альтернативных подходов к учету вмешивающихся факторов.	Исследователи могут не измерять или измерять неточно все значимые прогностические факторы. Неустановленные прогностические факторы могут приводить к систематическим ошибкам (т.е. за счет остаточных вмешивающихся факторов).
Анализ с учетом предпочтительного назначения определенного вмешательства (propensity analysis: (propensity-based matching, propensity score adjustment)	Метод, в соответствии с которым для каждого пациента помощью шкалы (propensity score) рассчитывают показатель предпочтительного применения вмешательства, который должен отражать в целом характеристики пациента. После этого либо выполняют регрессионный анализ с учетом оценке по шкале propensity, либо формируют пары со сходной оценкой по шкале propensity.	Позволяет улучшить результаты анализа при большом числе прогностических факторов в тех случаях, когда оценивается небольшое число клинических исходов. Может отчетливо продемонстрировать сбалансированность групп в диапазоне каждого из децилей оценки по шкале propensity, отражающей вероятность предпочтительного назначения исследуемого препарата.	Как и в случае выполнения стандартного регрессионного анализа с целью учета прогностических факторов, не позволяет учитывать остаточные вмешивающиеся факторы. На практике результаты часто сходны с таковыми при выполнении стандартного регрессионного анализа. Интерпретация менее интуитивно понятна.
Анализ инструментальных переменных (Instrumental variable analysis)	Цель метода состоит в установлении показателей, которые могут заменить рандомизацию за счет их связи с применением у пациента определенного вмешательства. Такой инструментальный показатель не должен иметь связи с каким-либо клиническим исходом, кроме изучаемого в данном исследовании.	Теоретически использование приемлемых инструментальных показателей позволяет учесть неизвестные (неустановленные) прогностические факторы, что невозможно достичь за счет выполнения регрессионного анализа или анализа propensity.	Приемлемые инструментальные показатели трудно найти, а их определение основано на сильных допущениях, некоторые из которых трудно проверить эмпирически, и они могут быть трудны для понимания.
Рандомизированные клинические исследования	Экспериментальное исследование, в котором пациенты рандомизированно распределяются либо в группу вмешательства, либо в группу контроля. Для дальнейшего уменьшения риска систематических ошибок применяют определенные методы (например, центральная рандомизация; использование слепого метода при распределении пациентов в группы; анализ данных исходя из допущения, что у всех пациентов применялось назначенное лечение; полное наблюдение за участниками).	Оптимальный подход к обеспечению сбалансированности группы вмешательства и контроля по прогностическим факторам. Обеспечивает сбалансированность групп как по известным (установленным), так и неизвестным (неустановленным) прогностическим факторам.	Не всегда возможно (например, для оценки эффектов в отдаленные сроки наблюдения, для выявления редко развивающихся побочных эффектов). Не всегда возможно по этическим соображениям.

мира. Наибольшее число членов ISPOR было из Северной Америки и Европы: их доля достигала 41% и 37% соответственно. Члены организации представлены многими заинтересованными сторонами, в частности, исследователями и представителями академической науки, экспертами по оценке экономической эффективности и представителями регуляторных органов, а также финансирующих организаций и разработчиков политики здравоохранения, поставщиков медицинских услуг и организаций по вовлечению пациентов в лечение. Уже более 20 лет ISPOR используется как независимый источник информации при принятии решения о недрении инноваций в медицину.

Исследования клинической практики по оценке эффектов прямых пероральных антикоагулянтов как источник доказательной информации

Опубликовано более большое число статей, посвященных оценке эффективности прямых оральных антикоагулянтов (ПОАК) у пациентов с фибрилляцией предсердий (ФП) в условиях клинической практики. Большая часть таких статей представляет собой отчет о ретроспективном анализе административных баз данных или баз данных страховых заявок. Следует отметить, что каждое из таких исследований имело определенные ограничения, обусловленные ретроспективным харак-

тером обсервационного исследования и использованием баз данных страховых заявок в качестве источника информации о пациентах. В связи с этим в рамках данной статьи обсуждение каждого из таких исследований не представляется обоснованным. В то же время в течение последних нескольких лет стали публиковаться статьи с отчетом о выполнении систематических обзоров и мета-анализов таких исследований. Анализ именно таких публикаций с результатами систематических обзоров и мета-анализов исследований по оценке эффективности и безопасности ПОАК в условиях клинической практики представляет больший интерес. Очевидно, что полученные в ходе выполнения систематических обзоров и мета-анализов данные о тенденциях позволяют точнее выявить тенденции особенностей применения ПОАК в условиях клинической практики, так как уменьшают возможные систематические ошибки отдельных исследований и возможное влияние конфликта интересов.

Для нахождения соответствующих статей был выполнен поиск в базе данных PubMed по ключевым словам и словосочетаниям: (apixaban AND dabigatran AND rivaroxaban) AND real clinical practice AND atrial fibrillation NOT suboptimal NOT cost-effectiveness. Выполняли поиск статей, опубликованных до 03 августа 2021 г. В результате было найдено 2 статьи с отчетом о выполнении систематического обзора и мета-анализа обсервационных исследований с оценкой эффективности применения ПОАК в клинической практике.

В ходе выполнения первого из них М. Proietti и соавт. [39] на основании мета-анализа 16 обсервационных исследований клинической практике (в целом в них было включено 170814 пациентов с ФП) установили, что применение апиксабана в стандартной дозе было более эффективно для снижения риска развития любого тромбоэмболического осложнения (отношение шансов 0,77 при 95% доверительном интервале [ДИ] от 0,64 до 0,93), но в отсутствие статистически значимых различий по риску развития инсульта. В то же время прием апиксабана был более эффективен по сравнению с приемом дабигатрана и ривароксабана для профилактики развития тромбоэмболических осложнений и инсульта. Риск развития тяжелых кровотечений при приеме апиксабана был статистически значимо меньше по сравнению с применением варфарина, дабигатрана и ривароксабана (снижение относительного риска на 38%, 35% и 46% соответственно). Риск развития внутричерепного кровоизлияния при приеме апиксабана был статистически значимо меньше по сравнению с приемом как варфарина, так и ривароксабана (снижение относительного риска на 46% и 54% соответственно), но не по сравнению с дабигатраном. Риск развития кровотечений из желудочно-кишечного тракта был высоко статистически значимо меньше при применении апиксабана по сравнению с любым из других антикоагулянтов (р<0,00001 для всех сравнений). Таким образом, результаты мета-анализа анализа исследований клинической практики в целом подтвердили результаты РКИ исследования ARISTOTLE, показавшие высокую безопасность применения апиксабана. Сходные результаты были получены и в ходе выполнения мета-анализа, выполненного С. Escobar и соавт. [40] и включавшего 27 исследований клинической практики по оценке эффектов ПОАК у пациентов с ФП. При применении апиксабана по сравнению с варфарина риск развития тяжелых кровотечений был статистически значимо ниже (отношение рисков 0,66 при 95% ДИ от 0,55 до 0,80).

Заключение

Результаты крупных РКИ, которые выполнены с соблюдением принятых методологических подходов, остаются золотым стандартом получения доказательной информации об эффективности применения вмешательств и становятся главным аргументом при решении вопроса об одобрении вмешательства регуляторными органами. В то же время как представители регуляторных органов, так и практикующие врачи проявляют интерес к результатам исследования клинической практики, которые позволяют ответить на вопрос о сохранении эффектов вмешательства, которые были установлены в ходе РКИ, но в популяции пациентов, характеристики которых могут отличаться от таковых у участников РКИ. При решении вопроса об использовании результатов исследования, представляющих собой анализ ДРКП, в качестве источника доказательной информации, который может влиять на принятие решения о тактике лечения, целесообразно оценивать качество таких исследований и убедиться в попытках авторов минимизировать вероятность систематических ошибок, ограничивающих обоснованность полученных данных. Следует учитывать, что ДРКП не соответствуют определению «большие данные», но представляют собой просто большое число данных. Разработанные критерии к оценке качества исследований клинической практики могут помочь такой оценке.

Отношения и Деятельность: Статья опубликована при финансовой поддержке компании Пфайзер. Компания Пфайзер не участвовала в получении данных и написании статьи. Мнение автора может не совпадать с мнением компании.

Relationships and Activities: The article was published with the financial support of the Pfizer company. Pfizer did not participate in the data acquisition and writing of the article. The opinion of the author may not coincide with the opinion of the company.

Assessing the Quality of Observational Studies Оценка качества наблюдательных исследований

References / Литература

- 1. U.S. Food and Drug Administration. Framework for FDA's real-world evidence programa (2018) [cited by Jun 30, 2021]. Available from: https://www.fda.gov/media/120060/download.
- 2. Berger ML, Doban V. Big data, advanced analytics and the future of comparative effectiveness research. J Comp Eff Res. 2014;3(2):167-76. DOI:10.2217/cer.14.2.
- 3. Okada M. Big data and real-world data-based medicine in the management of hypertension. Hypertens Res. 2021;44(2):147-53. DOI:10.1038/s41440-020-00580-3.
- Hay SI, George DB, Moyes CL, Brownstein JS. Big data opportunities for global infectious disease surveillance. PLoS Med. 2013;10(4):e1001413. DOI:10.1371/journal.pmed.1001413.
- Ola O, Sedig K. The challenge of big data in public health: an opportunity for visual analytics. Online J Public Health Inform. 2014;5(3):223. DOI:10.5210/ojphi.v5i3.4933.
- Fanaroff AC, Califf RM, Lopes RD. New Approaches to Conducting Randomized Controlled Trials. J Am Coll Cardiol 2020;75:556-9. DOI:10.1016/j.jacc.2019.11.043.
- Murray KW, Duggan A. Understanding Confounding in Research Pediatr Rev. 2010;31(3):124-6. DOI:10.1542/pir.31-3-124.
- Meuli L, Dick F. Understanding Confounding in Observational Studies. Eur J Vasc Endovasc Surg. 2018;55(5):737. DOI:10.1016/j.ejvs.2018.02.028.
- Fanaroff AC, Califf RM, Windecker S, et al. Levels of Evidence Supporting American College of Cardiology/American Heart Association and European Society of Cardiology Guidelines, 2008-2018. JAMA. 2019;321:1069-80. DOI:10.1001/jama.2019.1122.
- Gilyarevsky SR, Gavrilov DV, Gusev AV. Retrospective analysis of electronic health records of patients with heart failure: the first Russian experience. Russian Journal of Cardiology. 2021;26(5):4502 (In Russ.) [Гиляревский С.Р., Гаврилов Д.В., Гусев А.В. Результаты ретроспективного анализа записей электронных амбулаторных медицинских карт пациентов с хронической сердечной недостаточностью: первый российский опыт. Российский Кардиологический Журнал. 2021;26(5):4502]. DOI:10.15829/1560-4071-2021-4502.
- Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211-7. DOI:10.2147/JMDH.S104807.
- Matsumoto S, Fukui M, Hamaguchi M, et al. Is home blood pressure reporting in patients with type 2 diabetes reliable? Hypertens Res. 2014;37(8):741-5. DOI:10.1038/hr.2014.66.
- Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 2009;20(4):512-22. DOI:10.1097/EDE.0b013e3181a663cc.
- Cox DR, Kartsonaki C, Keogh RH. Big data: Some statistical issues. Stat Probab Lett. 2018;136:111-5. DOI:10.1016/j.spl.2018.02.015.
- Beaulieu-Jones BK, Finlayson SG, Yuan W, et al. Examining the Use of Real-World Evidence in the Regulatory Process. Clin Pharmacol Ther. 2020;107(4):843-52. DOI:10.1002/cpt.1658.
- Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363:k3532. DOI:10.1136/bmj.k3532.
- Benchimol El, Smeeth L, Guttmann A, et al.; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885. DOI:10.1371/journal.pmed.1001885.
- von Elm E, Altman DG, Egger M, et al.; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-7. DOI:10.1016/S0140-6736(07)61602-X.
- Hemkens LG, Benchimol EI, Langan SM, et al. The reporting of studies using routinely collected health data was often insufficient. J Clin Epidemiol. 2016;79:104-11. DOI:10.1016/j.jdinepi.2016.06.005.
- 20 Dreyer NA, Schneeweiss S, McNeil BJ, et al.; GRACE Initiative. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. Am J Manag Care. 2010;16(6):467-71.
- European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide
 on methodological standards in pharmacoepidemiology (2011) [cited by Jun 30, 2021]. Available
 from: www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml.
- 22. Blake KV, Devries CS, Arlett P, et al.; for the European Network of Centres for Pharmacoepidemiology Pharmacovigilance. Increasing scientific standards, independence and transparency in post-authorisation studies: the role of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Pharmacoepidemiol Drug Saf. 2012;21(7):690-6. DOI:10.1002/pds.3281.

- 23. Franklin JM, Schneeweiss S. When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? Clin Pharmacol Ther. 2017;102(6):924-33. DOI:10.1002/cpt.857.
- Schneeweiss S. Improving therapeutic effectiveness and safety through big healthcare data. Clin Pharmacol Ther. 2016;99(3):262-5. DOI:10.1002/cpt.316
- Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence What Is It and What Can It Tell Us? N Engl J Med. 2016;375(23):2293-7. DOI:10.1056/NEJMsb1609216.
- Guimarães PO, Krishnamoorthy A, Kaltenbach LA, et al. Accuracy of Medical Claims for Identifying Cardiovascular and Bleeding Events After Myocardial Infarction: A Secondary Analysis of the TRANS-LATE-ACS Study. JAMA Cardiol. 2017;2(7):750-7. DOI:10.1001/jamacardio.2017.1460.
- Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. In: The Cochrane Collaboration, ed. Cochrane Database of Systematic Reviews. Chichester, UK: John Wiley & Sons, Ltd; 2014 [cited by Jun 30, 2021]. Available from: http://doi.wiley.com/10.1002/14651858.MR000034.pub2. Accessed May 12, 2016
- Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. BMJ. 2016;352:i493. DOI:10.1136/bmj.i493.
- Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. Eur Heart J. 2012;33(15):1893-901. DOI:10.1093/eurhearti/ehs114.
- Stampfer MJ, Colditz GA, Willett WC, et al. Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the nurses' health study. N Engl J Med. 1991;325(11):756-62. DOI:10.1056/NEJM199109123251102.
- 31. Rossouw JE, Anderson GL, Prentice RL, et al.; Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA. 2002;288(3):321-33. DOI:10.1001/jama.288.3.321.
- Grodstein F, Manson JE, Stampfer MJ. Postmenopausal hormone use and secondary prevention of coronary events in the nurses' health study. a prospective, observational study. Ann Intern Med. 2001;135(1):1-8. DOI:10.7326/0003-4819-135-1-200107030-00003.
- Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology. 2008;19(6):766-79. DOI:10.1097/EDE.0b013e3181875e61.
- 34. Goodman SN, Schneeweiss S, Baiocchi M. Using Design Thinking to Differentiate Useful From Misleading Evidence in Observational Research. JAMA. 2017;317(7):705-7. DOI:10.1001/jama.2016.19970.
- Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. N Engl J Med. 2008;358(8):771-83. DOI:10.1056/NEJMoa0707571.
- Fergusson DA, Hébert PC, Mazer CD, et al.; BART Investigators. A comparison of aprotinin and lysine analogues in high-risk cardiac surgery. N Engl J Med. 2008;358(22):2319-31. DOI:10.1056/NEJ-Moa0802395.
- Agoritsas T, Merglen A, Shah ND, et al. Adjusted Analyses in Studies Addressing Therapy and Harm: Users' Guides to the Medical Literature. JAMA. 2017;317(7):748-59. DOI:10.1001/jama.2016. 20029.
- 38. ISPOR Membership profile [cited by May 12, 2021]. Available from: URL: https://www.ispor.org.
- Proietti M, Romanazzi I, Romiti GF, et al. Real-World Use of Apixaban for Stroke Prevention in Atrial Fibrillation: A Systematic Review and Meta-Analysis. Stroke. 2018;49(1):98-106. DOI:10.1161/STROKEAHA.117.018395.
- Escobar C, Martí-Almor J, Pérez Cabeza A, Martínez-Zapata MJ. Direct Oral Anticoagulants Versus Vitamin K Antagonists in Real-life Patients With Atrial Fibrillation. A Systematic Review and Meta-analysis. Rev Esp Cardiol (Engl Ed). 2019;72(4):305-16. DOI:10.1016/j.rec.2018.03.009.

Сведения об Авторе / About the Author Гиляревский Сергей Руджерович [Sergey R. Gilyarevsky]

Гиляревскии Сергеи гуджерович [sergey к. Gilyare ORCID 0000-0002-8505-1848