

УДК 519.23

ИСПОЛЬЗОВАНИЕ ПСЕВДОРАНДОМИЗАЦИИ (PROPENSITY SCORE MATCHING) ДЛЯ УСТРАНЕНИЯ СИСТЕМАТИЧЕСКИХ РАЗЛИЧИЙ МЕЖДУ СРАВНИВАЕМЫМИ ГРУППАМИ В ОБСЕРВАЦИОННЫХ ИССЛЕДОВАНИЯХ С ДИХОТОМИЧЕСКИМ ИСХОДОМ

© 2016 г. ¹⁻⁴А. М. Гржибовский, ⁵С. В. Иванов, ²М. А. Горбатова, ⁶А. А. Дюсупов

¹ Национальный институт общественного здравоохранения, г. Осло, Норвегия; ² Северный государственный медицинский университет, г. Архангельск; ³ Северо-Восточный федеральный университет, г. Якутск;

⁴ Международный казахско-турецкий университет, г. Туркестан, Казахстан; ⁵ Северо-Западный государственный медицинский университет им. И. И. Мечникова, г. Санкт-Петербург;

⁶ Государственный медицинский университет г. Семей, Казахстан

В настоящей статье представлен метод псевдорандомизации (propensity score matching – PSM) – эффективный статистический метод устранения влияния вмешивающихся факторов (конфаундеров), искажающих результаты обсервационных исследований при сравнении групп наблюдения. Метод PSM не уступает по эффективности регрессионному анализу и при этом не требует большого размера выборки. Рост популярности метода PSM среди исследователей подтверждается неуклонным ростом ежегодного числа исследований с его использованием в базе данных PubMed. В статье представлены теоретические основы метода и практическое его применение с использованием пакета программ для статистической обработки STATA 13. В практической части статьи подробно описаны пошаговые алгоритмы выполнения различных вариантов PSM (подбор пар и «взвешивание») для расчета отношения шансов с коррекцией на влияние конфаундеров

Ключевые слова: статистический анализ, обсервационные исследования, псевдорандомизация, propensity score, подбор пар, взвешивание, канфаундер, смещение результатов, ошибка отбора, бинарный исход, отношение шансов

PROPENSITY SCORE MATCHING AS A MODERN STATISTICAL METHOD FOR BIAS CONTROL IN OBSERVATIONAL STUDIES WITH BINARY OUTCOME

¹⁻⁴A. M. Grjibovski, ⁵S. V. Ivanov, ²M. A. Gorbatova, ⁶A. A. Dyussupov

¹ Norwegian Institute of Public Health, Oslo, Norway; ² Northern State Medical University, Arkhangelsk, Russia;

³North-Eastern Federal University, Yakutsk, Russia; ⁴ International Kazakh – Turkish University, Turkestan, Kazakhstan; ⁵ North-Western State Medical University, St. Petersburg, Russia;

⁶ Semey State Medical University, Semey, Kazakhstan

This article presents propensity score matching (PSM) - modern statistical method to control for confounding, which threatens the validity of associations in observational studies. The efficiency of PSM has been demonstrated in several international studies. The increasing use of PSM by the research community is reflected by steady growth of the number of publications with this method in the PubMed database over time. This article presents the theoretical basis of PSM and its practical application using Stata software. In the practical part of the article, detailed step by step algorithms of various PSM methods are presented allowing researchers to conduct statistical analysis of their own data and interpret results.

Keywords: statistical analysis, observational studies, pseudo-randomization, propensity score, matching, weighting, confounder, bias, selection bias

Библиографическая ссылка:

Гржибовский А. М., Иванов С. В., Горбатова М. А., Дюсупов А. А. Использование псевдорандомизации (propensity score matching) для устранения систематических различий сравниваемых групп в обсервационных исследованиях с дихотомическим исходом // Экология человека. 2016. № 5. С. 50–64.

Grjibovski A. M., Ivanov S. V., Gorbatova M. A., Dyussupov A. A. Propensity Score Matching as a Modern Statistical Method for Bias Reduction in Observational Studies with Binary Outcome. *Ekologiya cheloveka* [Human Ecology]. 2016, 5, pp. 50-64.

Как правило, основной целью исследований в медицине является корректная и точная оценка наличия и выраженности взаимосвязи между изучаемыми факторами воздействия и наблюдаемыми исходами. Изучаемыми факторами могут быть любые интересующие исследователя воздействия, такие как назначение определенных схем терапии, организационные мероприятия на уровне общественного здраво-

охранения, неблагоприятные факторы окружающей среды и др., а исходами могут быть, например, факт развития заболевания, смерть, развитие осложнения, ключевые количественные характеристики патофизиологического процесса и т.п.

В процессе планирования и проведения исследований следует принимать во внимание тот факт, что изучаемый исход подвержен как непосредственно

изолированному влиянию изучаемого в ходе исследования фактора, так и другим внешним и внутренним факторам (конфаундерам), которые могут оказывать сравнимое с изучаемым фактором (или даже большее) влияние на исход [13]. Следствием действия конфаундеров является искажение (смещение) результатов исследования, вплоть до получения результатов, противоположных реальному положению вещей и здравому смыслу.

Принципиальная схема связей между фактором (Е), исходом (О) и конфаундерами (С) представлена на рис. 1.

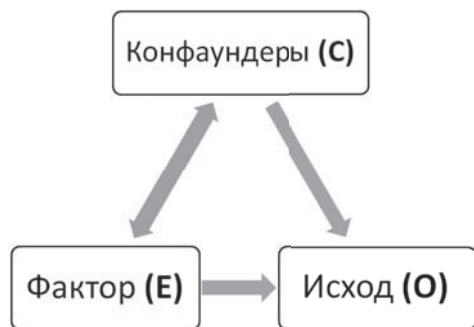


Рис. 1. Связи между фактором, исходом и конфаундерами

Именно поэтому «золотым стандартом» оценки эффекта изучаемого фактора являются рандомизированные контролируемые исследования (РКИ), так как в случае корректной рандомизации обеспечивается сбалансированное распределение известных и неизвестных конфаундеров между основной группой и группой контроля, что позволяет изолированно оценить эффект изучаемого фактора (например, нового лекарственного препарата). В РКИ устраняется обходной путь от Е до О через С, и путь от Е до О становится единственным. Таким образом, в результате корректного проведения РКИ влияние фактора на исход может быть оценено непосредственно путем сравнения исходов в основной и контрольной группах [4].

Несмотря на неоспоримое преимущество РКИ перед другими видами эпидемиологических исследований, они имеют существенные ограничения, так как с их помощью можно изучать достаточно ограниченное количество интересующих исследователей эффектов, в том числе вследствие невозможности организовать процедуру рандомизации или неэтичности исследований, направленных на изучение лечебного воздействия. Именно по этой причине одним из важнейших инструментов оценки связи между фактором и исходом в медицине остаются наблюдательные исследования [1, 2, 4].

Сложность корректной оценки мер эффекта изучаемого фактора, рассчитанных на основании данных наблюдательных исследований, связана с наличием конфаундеров, которые потенциально могут оказывать (и оказывают) влияние не только на исход (путь от С до О), но и на саму вероятность объекта исследования

оказаться под воздействием изучаемого фактора (путь от С до Е). Например, в случае изучения влияния физической активности пациентов на выраженность артериальной гипертензии таким конфаундером может быть возраст пациентов, так как уровень артериального давления (исход) с возрастом может увеличиваться, а чем больше возраст пациентов, тем менее вероятной представляется их значительная физическая активность (изучаемый фактор). Таким образом, игнорирование конфаундеров приводит к смещению результатов исследования — появлению существенных различий между рассчитанным и истинными значениями эффекта.

В ходе анализа данных наблюдательных исследований всегда имеется большая вероятность включить в анализ такие группы сравнения, в которых основная группа наблюдения имеет некие существенные систематические отличия от группы контроля, что даже в случае корректных расчетов меры эффекта фактора не позволит генерализовать результаты исследования вследствие ошибки отбора («selection bias»).

Таким образом, в случае наблюдательных исследований необходим такой способ включения наблюдений в анализ (не в исследование), который мог бы обеспечить максимальную «похожесть» основной группы и группы сравнения по имеющимся конфаундерам. Одним из таких способов является метод «Propensity score matching» (PSM), который был предложен Р. R. Rosenbaum и D. B. Rubin в 1983 г. [15]. Несмотря на то, что данный метод был предложен более 30 лет назад, его широкое внедрение в исследовательский процесс в медицине, эконометрике, психологии и социологии фактически активно началось только после 2000 г. и в настоящее время его использует все большее число исследователей по всему миру. Данный факт подтверждается неуклонным ростом числа публикаций, в которых упоминается или используется метод PSM, в базе данных национального центра биотехнологической информации (NCBI) PubMed (рис. 2).

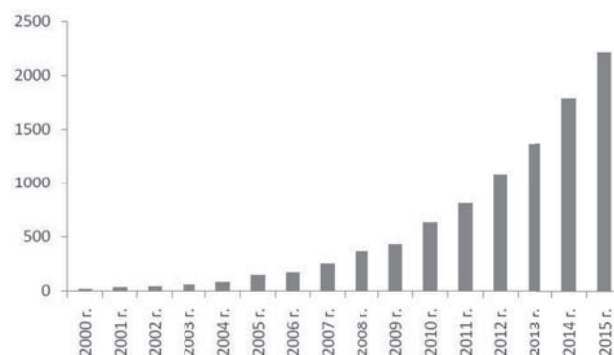


Рис. 2. Количество результатов поиска публикаций по ключевым словам «propensity score» в базе данных PubMed

Метод PSM основан на использовании значений индекса соответствия (propensity score — PS) — условной вероятности попадания каждого объекта исследования в основную или контрольную группу наблюдения на основании набора его характеристик.

То есть PS показывает вероятность для каждого наблюдения попасть под действие изучаемого фактора. В обсервационных исследованиях, в отличие от экспериментальных, исследователь не может произвольно каким-либо образом распределить действие фактора между группами наблюдения, и такое распределение происходит самопроизвольно, вследствие как случайности, так и набора неких характеристик объекта (конфаундеров), которые и определяют значение PS для каждого наблюдения.

Так как PS определяет условную вероятность наступления события, математические значения PS находятся в пределах от 0 до 1. Подробное математическое обоснование и описание метода PSM приведены в соответствующей литературе [8, 10, 12, 15].

Имея значения условной вероятности попадания объекта исследования под действие изучаемого фактора, можно подобрать группу сравнения таким образом, чтобы данная вероятность была сбалансирована между основной и контрольной группами. Таким образом, компенсируется неравномерное распределение конфаундеров между основной группой и группой сравнения, то есть проводится процедура, преследующая те же цели, что и рандомизации в РКИ, но только проводимая на этапе обработки, а не сбора исследовательских данных.

Так как значения PS рассчитываются на основании определенного набора характеристик объектов исследования, потенциально оказывающих влияние как на исход, так и вероятность попадания наблюдения в основную группу, данный метод позволяет свести широкий набор характеристик каждого наблюдения (потенциальных конфаундеров) к единому вариационному ряду значений PS [8, 10, 11].

Алгоритм PSM включает в себя следующие основные этапы:

1. Отбор переменных-ковариат, кодирующих конфаундеры, для включения в анализ.
2. Расчет значений PS.
3. Проверка баланса распределения значений PS и средних значений ковариат между основной и контрольной группами.
4. Выбор одного из методов подбора пар или «взвешивания» и расчет мер эффекта воздействия фактора.
5. Оценка степени уменьшения дисбаланса ковариат после применения выбранного метода подбора пары или «взвешивания».
6. Повтор анализа с использованием других методов подбора пар или «взвешивания».
7. Выбор для расчета результирующей меры эффекта того метода, который обеспечил в данном случае наилучший баланс ковариат между основной группой и группой сравнения.

Переменные, кодирующие фактор и исход, должны быть определены еще на этапе планирования исследования [1, 2, 4]. Переменная изучаемого фактора должна быть бинарной (есть фактор / нет фактора), переменная исхода должна быть количественной или

качественной. Следует отметить, что в зарубежных источниках кодирующая фактор переменная традиционно называется словом «treatment» — «лечение», в том числе и в эконометрике, психологических и социологических науках, хотя спектр факторов, названный этим словом, разумеется, гораздо шире, чем лечебное воздействие.

Отбор ковариат для включения в анализ, согласно самой концепции PSM, должен быть основан на теоретических представлениях об их потенциальной взаимосвязи и с изучаемым фактором, и с исходом [6]. На этапе планирования исследования требуется детализированная регистрация данных конфаундеров при сборе исследовательских данных.

Но с практической точки зрения на этапе сбора данных возможен качественный и количественный учет далеко не всех вероятных конфаундеров, поэтому желательно регистрировать и в дальнейшем включать в анализ такие характеристики объектов исследования, которые могут потенциально коррелировать с неучтенными конфаундерами, чтобы максимально уменьшить долю создаваемого ими смещения результатов анализа. При этом в анализ не должны включаться переменные, которые сами напрямую зависят от факта попадания объекта под действие фактора, и те, которые позволяют точно предсказать факт отнесения объекта исследования к основной или контрольной группе. Следует отметить, что многие вопросы отбора конфаундеров в PSM не имеют однозначных ответов и остаются объектом обсуждения в литературе [6, 11].

Примерами подобных конфаундеров могут служить демографические характеристики участников исследования (пол, возраст, национальная принадлежность и т. п.), характеристика коморбидных состояний (тяжесть заболевания, наличие сопутствующих заболеваний, длительность анамнеза, частота госпитализаций за определенный период времени и т. п.), локальные особенности (регион проживания, особенности климата и т. п.), а также другие характеристики объектов исследования, удовлетворяющие вышеперечисленным условиям.

Расчет значений PS производится с помощью логистического регрессионного анализа (или с использованием пробит-модели), в котором в качестве зависимой переменной указывается факт принадлежности к основной группе (группе «лечения»), закодированный в виде бинарной переменной (да / нет), а кодирующие конфаундеры ковариаты являются независимыми переменными. На этапе расчета значений PS переменная исхода не используется.

Таким образом, в анализ включается следующий ряд переменных:

1. Переменная «лечение» (действующий фактор), которая имеет бинарный характер, то есть кодируется значениями, имеющими только два альтернативных варианта (например, терапия 1 / терапия 2, наличие фактора / отсутствие фактора и т. п.).
2. Переменная исхода, которая может иметь как количественный, так и бинарный характер. Следует

отметить, что в настоящей публикации рассматриваются случаи бинарного исхода как наиболее часто оцениваемой конечной точки исследования. С практической точки зрения использование бинарной переменной исхода позволяет оценивать многие истинные клинически значимые исходы исследования (выжил / умер, имеется осложнение / отсутствует осложнение и т. п.), в то время как количественные переменные в большинстве случаев позволяют оценивать суррогатные исходы [4].

3. Переменные-конфаундеры, которые могут иметь количественный, номинальный, порядковый, бинарный характер (например, пол, возраст, степень тяжести заболевания, наличие сопутствующей патологии и т. п.).

Пример распределения значений PS в основной группе и группе сравнения представлен на рис. 3 [6, 14, 16]. Закономерно, что рассчитанные значения условной вероятности попадания наблюдения в основную группу могут распределиться таким образом, что некая часть наблюдений будет иметь набор характеристик, создающий очень высокую вероятность подвергнуться действию фактора, а для другой части, напротив, попадание под действие фактора будет крайне маловероятно. Таким образом, для сбалансированного сравнения эффекта изучаемого фактора наиболее подходящими будут наблюдения, попавшие в зону «перекреста» областей значений PS («common support»).



Рис. 3. Пример распределения значений PS в основной группе и в группе сравнения

Баланс распределения значений PS между основной и контрольной группами проводится путем разделения выборки наблюдений на страты таким образом, чтобы средние значения PS в пределах каждой страты в основной группе и группе сравнения статистически значимо не различались. После определения страт проводится проверка баланса всех ковариат в пределах каждого блока.

Далее используется один из следующих основных методов анализа [6, 8, 11]:

– Стратификация (субклассификация) объектов исследования на основании значений PS.

– Подбор к каждому наблюдению основной группы одной или нескольких пар — «ближайших соседей»

– наблюдений с наиболее близким значением PS. В этом случае в анализ включаются все наблюдения основной группы и удаляется часть наблюдений группы сравнения, которые оказались не соответствующими ни одному из наблюдений основной группы. Недостатком данного метода является потеря части информации о выборке по причине удаления части наблюдений и тот факт, что «ближайшее» значение PS некоторых наблюдений из группы контроля может сильно отличаться от значений PS соответствующих наблюдений основной группы.

– Подбор к каждому наблюдению основной группы одной или нескольких пар, значения PS которых укладываются в пределы заданного числового отрезка. Таким образом исключается вероятность включить в анализ контрольное наблюдение, достаточно сильно удаленное по значению PS от соответствующего наблюдения основной группы.

– «Взвешивание» Кернела, при котором каждому наблюдению основной группы присваивается весовой коэффициент, равный 1, а все наблюдения группы сравнения ранжируются таким образом, чтобы более близкие к объектам основной группы наблюдения контрольной группы получали большие «веса» по сравнению с теми, которые оказались более удаленными.

Вышеперечисленные методы используются на первом этапе PSM и необходимы для балансировки конфаундеров между основной и контрольной группой.

Непосредственный расчет меры эффекта действующего фактора для бинарного исхода производится на следующем этапе обработки данных. Существуют следующие варианты использования PS для коррекции результатов анализа:

1. Включение в логистическую регрессионную модель значений PS в качестве единственной переменной-конфаундера наряду с бинарными переменными исхода (зависимая переменная) и «лечение».

2. Логистический регрессионный анализ на базе редуцированной выборки, включающей в себя только те наблюдения основной и контрольной групп, которые составили пары по результатам обработки данных на предыдущем этапе анализа.

3. Логистический регрессионный анализ на базе «взвешенной» выборки.

4. Метод Мантеля — Хензеля с использованием страт, сформированных на предыдущем этапе анализа.

Метод PSM является методом коррекции влияния конфаундеров на результаты исследования, альтернативным по отношению к широко используемому для этой цели регрессионному анализу, поэтому сравнение эффективности данных двух методов представляет большой интерес и широко обсуждается в литературе [6, 11].

В первую очередь следует отметить, что метод PSM более точен по сравнению с регрессионным анализом в случае наличия относительно небольшого числа наблюдений на фоне большого числа конфаундеров, которые требуется включить в анализ. Эмпирически выяснено, что при наличии 7 и менее наблюдений на

каждый конфаундер метод PSM обеспечивает более точную (менее смещенную) оценку меры эффекта по сравнению с регрессионным анализом [9]. Другим существенным преимуществом метода PSM перед регрессионным анализом служит то, что вследствие определенных допущений регрессионных моделей они не всегда могут корректно описать фактические данные.

Тем не менее не следует забывать, что регрессионный анализ является общепринятым статистическим методом коррекции результатов исследования, имеет колоссальный накопленный опыт использования и, в отличие от PSM, позволяет рассчитать меру эффекта изолированного для каждого включенного в анализ конфаундера, что может быть одной из основных задач исследования.

Следует учесть, что оба метода предназначены для коррекции влияния учтенных конфаундеров, в то время как смещение результатов может возникать вследствие неучтенных конфаундеров, и оценка степени смещения результатов в таком случае оказывается весьма затруднительной.

Как и в случае любого многоэтапного метода статистической обработки данных, использование метода PSM требует наличия специализированного статистического программного обеспечения, к которому относятся программные пакеты SAS, SPSS, R и STATA.

SAS и R являются программами, изначально требующими серьезного и углубленного освоения, и по этой причине их использование на этапе знакомства с новым методом представляется нецелесообразным. Программа SPSS имеет модуль PSM только в последней версии, а предыдущие версии программы требуют установки дополнительных пакетов совместимости.

В данной статье рассматривается профессиональный статистический программный пакет STATA 13.0 (STATA Corp, TX, USA) — удобный инструмент для выполнения статистического анализа данных, доступный для достаточно быстрого освоения [3]. Официальный сайт разработчика программы — www.stata.com. Программа STATA позволяет программировать последовательность команд обработки данных и детализировать весь процесс анализа, который проводится путем ввода специализированных команд, имеющих определенный синтаксис [5]. Анализ может проводиться и с использованием меню модуля «Statistics», который тем не менее имеет ограниченные функциональные возможности по сравнению с командным способом проведения анализа.

В настоящей статье будет рассмотрен синтаксис команд STATA и пошагово разобран алгоритм обработки данных с использованием различных вариаций метода PSM [11], а также интерпретация полученных результатов.

Для удобства восприятия в настоящей статье команды STATA будут записываться традиционным для данной программы образом:

— Жирным шрифтом — непосредственно команды обработки данных (например, **pscore**).

— Курсивом — вариационные ряды, представленные в исходной базе данных (например, *complicat*).

— Обычным шрифтом — названия новых вариационных рядов, создаваемых в процессе анализа (например, *mypsore*).

Синтаксис команд включает названия переменных из рассматриваемой базы данных, так что строку команды можно копировать из текста статьи непосредственно в поле «Command» программы STATA и запускать ее выполнение клавишей «Enter». Следует учесть, что в программе STATA в качестве десятичного разделителя используется не запятая, а точка, что важно при вводе значений в базу данных (в противном случае такие переменные будут восприниматься программой как качественные). Также отметим, что поскольку программа практически не имеет возможности работать с кириллическими символами, для открытия и сохранения файлов рекомендуется использовать такие места хранения, которые в наименовании пути к файлу не имели бы русскоязычных названий папок (например, сохранять все файлы в какой-нибудь корневой каталог).

Для начала работы необходимо загрузить из интернета дополнительные модули программы, предназначенные для PSM. Для этого введем в поле «Command» рабочего поля программы STATA команду `findit pscore`.

Программа откроет новое окно, в котором выберем модуль `st0026_2` и нажмем «click here to install» для установки данного модуля в имеющийся программный пакет STATA 13.

Заметим, что в дальнейшем, если программа показывает, что введенная команда не распознана (сообщение «unrecognized command»), необходимо с помощью команды `findit` найти соответствующий модуль и установить его, следуя инструкциям системы, подобно тому, как это было продемонстрировано в отношении команды `pscore` (может потребоваться в отношении команды `psmatch2` и др.).

Для наглядного примера выполнения PSM рассмотрим базу данных одного из исследований, проведенного казахстанскими учеными, целью которого была сравнительная оценка эффективности двух альтернативных видов оперативных вмешательств при аневризме брюшного отдела аорты. Для загрузки базы данных в программу необходимо прежде всего загрузить ее с сайта журнала (файл `DataSTATA.dta`) и открыть ее через меню «File» — «Open».

В базе данных представлены следующие переменные:

— *group* (группа наблюдения): кодирует «лечение» и имеет бинарный характер (один из двух альтернативных вариантов оперативного вмешательства). Следует отметить, что в качестве основной группы целесообразно использовать ту группу, в которой имеется меньшее число наблюдений. Основная группа всегда кодируется цифрой «1», контрольная — цифрой «0». В нашем случае кодировка выбрана так, чтобы в основной группе «1» (тип операции 1) было меньше наблюдений, чем в контрольной «0» (тип операции

2), для расширения возможностей по подбору пар из группы сравнения к наблюдениям основной группы.

– *complicat* (наличие осложнений): кодирует исход и имеет значения «1» (развилось осложнение) и «0» (осложнений не было).

– *gender, age, length, diam* – соответственно пол и возраст пациентов, длина и диаметр шейки аневризмы. В нашем случае для уменьшения громоздкости примера были выбраны только эти ковариаты, из которых первая имеет качественный (бинарный), а остальные – количественный тип.

Как было сказано выше, ключевым вопросом отбора переменных в анализ является определение переменных-ковариат (конфаундеров), которые потенциально оказывают влияние как на исход, так и на вероятность участника исследования подвергнуться действию изучаемого фактора. Разумеется, в реальных исследованиях учитывается гораздо более широкий набор конфаундеров, но в нашем примере будут использованы только три из них для уменьшения громоздкости изложения материала в статье.

На предварительном этапе обработки данных с помощью программы STATA 13 целесообразно оценить, насколько отличаются группы по основным ковариатам. Так, для оценки распределения количественной переменной *age* между сравниваемыми группами воспользуемся командой:

by group, sort : summarize age, detail

В результате будет представлена описательная статистика возраста пациентов в разбивке по группам: процентилю, среднее значение, стандартное отклонение и др. (результаты не приводятся). Подставляя в синтаксис команды переменные *length* и *diam*, можно оценить их распределение между группами.

Для оценки распределения между группами качественной переменной *gender* воспользуемся командой:

tab group gender

Данная команда выводит четырехпольную таблицу, в которой программа представляет количество мужчин и женщин в каждой группе.

В нашем примере имеются различия между сравниваемыми группами по всем ковариатам, в чем читатель может убедиться самостоятельно с помощью использования вышеуказанных команд.

	group	complicat	gender	age	length	diam	u	mypscore	myblock	comsup
1	1	0	0	54	25	15	.3179207	.27791215	2	1
2	0	0	0	59	22	15	.1541927	.34152019	2	1
3	0	1	0	77	60	10	.1441176	.89349632	5	1
4	1	1	0	55	36	19	.630199	.58158674	3	1
5	1	0	0	67	30	15	.9784231	.51154632	3	1
6	1	0	0	66	22	16	.6952117	.41703612	3	1
7	1	0	0	72	19	15	.0119713	.55073047	3	1
8	1	0	0	79	30	22	.9206055	.80203126	5	1
9	0	0	0	73	60	35	.2878177	.33083349	2	1
10	0	0	0	65	25	25	.7725696	.45454615	3	1
11	0	0	0	71	22	25	.7503774	.54642224	3	1
12	0	0	0	73	20	25	.7589197	.57235417	3	1
13	1	0	0	75	12	21	.015148	.72535048	4	1

Рис. 5. Переменные базы данных DataSTATA.dta

1. Установка случайного порядка наблюдений

Данное действие очень важно для правильного расчета PS, так как систематизированное расположение наблюдений (например, сначала наблюдения основной группы, затем – контрольной) может исказить результаты расчета PS. С этой целью используются команды, которые необходимо построчно ввести в поле «Command» и каждую запустить клавишей «Enter»:

```
set seed 123456
gen u=uniform()
sort u
```

2. Расчет значений PS

Для расчета значения PS для каждого наблюдения и ряда других промежуточных значений введем в поле «Command» следующую команду (рис. 4):

```
pscore group gender age length diam,
pscore(mypscore) blockid(myblock) logit
detail comsup
```



Рис. 4. Ввод команды в поле «Command» для расчета значений PS

Следует заметить, что после команды **pscore** сразу же следует название переменной, кодирующей изучаемый фактор («лечение»), после чего последовательно указываются названия ковариат. Команда создает новую переменную со значениями PS, имеющую любое удобное название, которое мы указываем в скобках (в нашем случае выберем название *mypscore*). Команда **logit** указывает на то, что мы используем логистическую регрессию для расчета значений PS (если ее удалить, то будет использована пробит-модель, что также вполне допустимо), команда **detail** позволяет детализировать все промежуточные расчеты, что может оказаться важным на последующих этапах анализа, а команда **comsup** фиксирует зону «перекреста» значений PS между основной группой и группой сравнения («common support»).

В результате выполнения команды **pscore** программа создаст ряд новых переменных в дополнение к уже имеющимся в базе данных: рассчитает значение

group	Freq.	Percent	Cum.
0	118	61.14	61.14
1	75	38.86	100.00
Total	193	100.00	

Рис. 6. Таблица распределения наблюдений в зависимости от значения переменной «лечения» (*group*)

PS для каждого наблюдения (переменная *mypscore*), укажет номер блока, в который объединены значения PS (переменная *myblock*) и факт соответствия каждого наблюдения области «перекреста» (переменная *comsup*). Созданные переменные можно увидеть, если войти в меню «Data» – «Data Editor» – «Data Editor (Edit)» (рис. 5).

В процессе выполнения анализа программа последовательно выводит результаты расчетов, которые мы будем рассматривать, акцентируя внимание на ключевых моментах.

На рис. 6 представлена сводная таблица, в которой указано количество наблюдений, соответствующих значению переменной «лечение» – *group*: в базе данных представлены общие сведения о 75 наблюдениях основной группы и 118 наблюдениях группы контроля.

Заметим, что в итоговых таблицах программа точкой отделяет десятичные знаки, а запятой разделяет разряды чисел.

Результаты логистического регрессионного анализа, использованного для расчета значений PS, представлены на рис. 7. Программа указывает псевдокоэффициент детерминации («Pseudo R2»), коэффициенты для каждой ковариаты («Coef.») и константу («_cons»). В нижней части рисунка указано, какой промежуток значений PS соответствует области «перекреста» (от 0,055 до 0,937).

Программа также представляет описательную статистику вариационного ряда значений PS в виде процентилей (рис. 8), причем учитываются только значения PS, ограниченные областью «перекреста». Из представленной таблицы видно, что медиана значений PS равна 0,405 (процентиль «50 %»), среднее значение – 0,419 («Mean»). Следует отметить, что в нашем примере в зону «перекреста» попадают 175 случаев из имеющихся в базе данных 192 наблюдений («Obs»).

Для наглядного представления о распределении значений PS в основной группе и в группе сравнения сформируем гистограмму распределения (рис. 9) с помощью соответствующей команды:

```
psgraph, bin(15) treated(group)
pscore(mypscore)
```

В данной команде значение «15» соответствует количеству столбиков гистограммы и может быть произвольно изменено (например, можно указать любое целое число в зависимости от объема выборки и наглядности полученной гистограммы).

```
Logistic regression                                Number of obs =      192
LR chi2(4) =      54.13
Prob > chi2 =      0.0000
Pseudo R2 =      0.2115

Log likelihood = -100.93224
```

group	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gender	-.4826679	.5331808	-0.91	0.365	-1.527683	.5623472
age	.0702294	.0247622	2.84	0.005	.0216963	.1187624
length	.0176107	.0102512	1.72	0.086	-.0024812	.0377026
diam	-.1700309	.0326741	-5.20	0.000	-.2340709	-.1059908
_cons	-.9367235	1.847591	-0.51	0.612	-4.557936	2.684488

Note: the common support option has been selected
The region of common support is [.05471478, .93658473]

Рис. 7. Результаты логистической регрессии с целью расчета значений PS

Percentiles		Smallest		
1%	.0570473	.0547148		
5%	.0707002	.0570473		
10%	.0929894	.0573738	Obs	175
25%	.2423506	.0580487	Sum of Wgt.	175
50%	.4053056		Mean	.4191605
		Largest	Std. Dev.	.2306478
75%	.5975322	.8582808	Variance	.0531984
90%	.7568183	.8636621	Skewness	.1862349
95%	.8020313	.8934963	Kurtosis	2.026603
99%	.8934963	.9365847		

Рис. 8. Описательная статистика значений PS для области «перекреста»

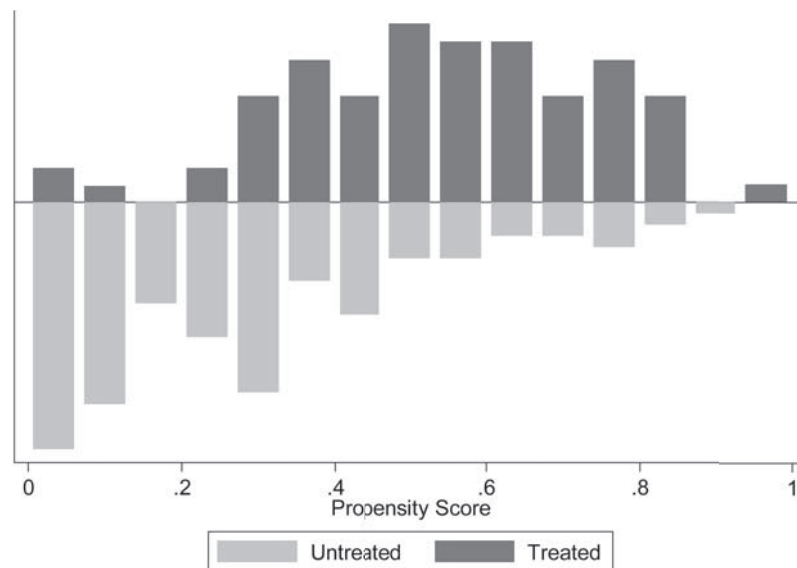


Рис. 9. Гистограмма распределения значений PS в основной группе («Treated») и группе сравнения («Untreated»): по оси абсцисс – значение PS, по оси ординат – количество наблюдений

На представленной двойной гистограмме видно, что распределение значений PS в основной и контрольной группах существенно различается, то есть имеются существенные различия между группами по распределению конфаундеров (пол, возраст, длина и диаметр шейки аневризмы).

3. Формирование страт (блоков) PS и проверка баланса (равновесия) значений PS между основной группой и группой сравнения в пределах каждого блока PS

Далее команда `pscore` разделяет выборку на наблюдений на блоки таким образом, чтобы в границах блоков значения PS в основной группе и группе сравнения статистически значимо не различались (рис. 10).

Blocks of the pscore for treatment group	group		Total
	0	1	
1	32	3	35
2	36	16	52
3	20	25	45
4	10	23	33
5	3	7	10
Total	101	74	175

Рис. 10. Результаты деления выборки на блоки

Затем программа проверяет наличие статистических различий между основной группой и группой сравнения в пределах каждого блока и, если они имеются, делит блок надвое и проверяет наличие статистически значимых различий снова уже в пределах новых блоков, полученных в результате деления исходного. Данный

процесс продолжается до тех пор, пока значения PS в пределах каждого их сформированных блоков не перестанут статистически значимо различаться при сравнении основной и контрольной групп.

Рассмотрим в качестве примера блок № 4 (рис. 11), который состоит из 23 наблюдений основной группы и 10 наблюдений группы сравнения.

В таблице представлены среднее значение PS в основной и контрольной группах в пределах блока («Mean»), его стандартная ошибка («Std. Err.») и стандартное отклонение (Std. Dev.). Сравнение групп показало отсутствие статистически значимых различий значений PS между основной и контрольной группами («Pr(T > t) = 0.9360», о чем программа выдала соответствующее сообщение («The mean propensity score is not different for treated and controls in block 4»).

Если в каком-либо из блоков средние значения PS статистически значимо различаются (значение «Pr(T > t)» составляет менее 0,05), программа представляет сообщение «The mean propensity score is different for treated and controls in block»), после чего разделяет данный блок на две части, каждую из которых проверяет в процессе последующего анализа (сообщение «Split the block and retest»), повторяя данную процедуру до тех пор, пока во всех получившихся блоках не будут отсутствовать статистически значимые различия PS между основной и контрольной группами. В нашем примере не потребовалось разделение ни одного из блоков и итоговое количество блоков осталось равным 5 (см. рис. 10).

Среди представленных результатов программа также напоминает, что в анализ включены наблюдения, значения которых соответствуют зоне «перекреста» («Note: the common support option has been selected»).

4. Проверка всех включенный в PSM ковариат (конфаундеров) на предмет баланса между основной группой и группой сравнения в блоках PS

Test for block 4

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	10	.6957241	.0169033	.0534528	.6574863	.7339619
1	23	.6975658	.0130332	.0625052	.6705365	.7245951
combined	33	.6970077	.0102843	.059079	.6760592	.7179562
diff		-.0018417	.0227339		-.0482078	.0445244

diff = mean(0) - mean(1) t = -0.0810
 Ho: diff = 0 degrees of freedom = 31

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
 Pr(T < t) = 0.4680 Pr(|T| > |t|) = 0.9360 Pr(T > t) = 0.5320

The mean propensity score is not different for treated and controls in block 4

Рис. 11. Результаты проверки блока № 4 на предмет наличия статистически значимых различий значений PS между основной группой и группой сравнения

Далее программа проводит сравнение значений ковариат в основной и контрольной группах в рамках каждого из блоков. Сравнение осуществляется таким же образом, как на предыдущем этапе, но сравниваемыми вариационными рядами являются уже не значения переменной *myscore*, а средние значения ковариат – *gender*, *age*, *length* и *diam*.

Пример сравнения ковариаты *age* в основной и контрольной группах в пределах блока № 2 представлен на рис. 12. В данной таблице программа указывает среднее значение вариационного ряда данной переменной в основной и контрольной группах («Mean»), стандартную ошибку («Std. Err.») и стандартное отклонение (Std. Dev.). Сравнение групп показало отсутствие статистически значимых различий значений ковариаты *age* между основной и контрольной группами («Pr(T > t) = 0.9975», сообщение «Variable age is balanced in block 2»).

В случае отсутствия баланса определенной ковариаты в рамках отдельного блока программа информирует об этом исследователя соответствующим сообщением: например, сообщение программы «Variable length is not balanced in block 3» говорит об

отсутствии баланса ковариаты *length* в блоке № 3. В результате анализа выявлен дисбаланс ковариаты *length* в блоках № 3 и 4, в то время как все остальные ковариаты оказались сбалансированными во всех блоках. Программа информирует о наличии дисбаланса в некоторых блоках сообщением «The balancing property is not satisfied. Try a different specification of the propensity score».

Итак, баланс всех ковариат во всех блоках не достигнут, и возникает вопрос, как поступать в таком случае? В целом, согласно опыту использования PSM, идеального баланса всех ковариат в каждом блоке, как правило, достигнуть не удается. Следует обдуманно взвесить потенциальное влияние несбалансированного конфаундера на исход и вероятность участника исследования оказаться под влиянием изучаемого фактора: если на основании теоретических данных потенциальное влияние данного конфаундера отно-

шим сообщением: например, сообщение программы «Variable length is not balanced in block 3» говорит об отсутствии баланса ковариаты *length* в блоке № 3. В результате анализа выявлен дисбаланс ковариаты *length* в блоках № 3 и 4, в то время как все остальные ковариаты оказались сбалансированными во всех блоках. Программа информирует о наличии дисбаланса в некоторых блоках сообщением «The balancing property is not satisfied. Try a different specification of the propensity score».

Итак, баланс всех ковариат во всех блоках не достигнут, и возникает вопрос, как поступать в таком случае? В целом, согласно опыту использования PSM, идеального баланса всех ковариат в каждом блоке, как правило, достигнуть не удается. Следует обдуманно взвесить потенциальное влияние несбалансированного конфаундера на исход и вероятность участника исследования оказаться под влиянием изучаемого фактора: если на основании теоретических данных потенциальное влияние данного конфаундера отно-

Testing the balancing property for variable age in block 2

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	36	63.94444	1.213642	7.281854	61.48062	66.40827
1	16	63.9375	1.806167	7.224668	60.08775	67.78725
combined	52	63.94231	.9975133	7.193171	61.93972	65.9449
diff		.0069444	2.182784		-4.377307	4.391196

diff = mean(0) - mean(1) t = 0.0032
 Ho: diff = 0 degrees of freedom = 50

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
 Pr(T < t) = 0.5013 Pr(|T| > |t|) = 0.9975 Pr(T > t) = 0.4987

Variable age is balanced in block 2

Рис. 12. Результаты проверки блока № 2 на предмет наличия статистически значимых различий значений ковариаты *age* между основной группой и группой сравнения

сительно невелико, то небольшой дисбаланс таких ковариат допустим. В ином случае существует ряд способов достижения баланса:

- Использование не логит-, а пробит-модели расчета PS (для этого потребуется убрать опцию **logit** из команды **pscore**).
- Удаление из анализа переменных, связанных с уже включенными в анализ ковариатами.
- Включение в анализ одной или нескольких дополнительных ковариат.
- Преобразование несбалансированных ковариат (например, преобразование уровня гемоглобина в ряд категорий: «анемии нет», «анемия легкой степени», «анемия средней степени тяжести» и т. д.).

После выполнения вышеуказанных процедур потребуется повторить вышеперечисленные этапы анализа и снова оценить степень балансировки блоков PS и ковариат.

Заметим, что прежде чем повторять анализ, необходимо удалить созданные программой переменные:

```
drop mypscore myblock comsup
```

Далее описываются методы, с помощью которых производится окончательная балансировка конфаундеров между основной и контрольной группами.

5. Подбор метода, позволяющего наиболее адекватно сбалансировать основную и контрольную группы по всем ковариатам

5.1. Метод подбора пар 1 : 1 с помощью поиска «ближайшего соседа»

В данном случае каждому наблюдению основной группы подбирается «пара» из группы контроля, имеющая наиболее близкое значение PS.

Команда **psmatch2** использует рассчитанные ранее значения PS:

```
psmatch2 group, pscore(mypscore) noreplacement
```

Исполнение такого синтаксиса команды не приводит к появлению изменений на экране, но она создает ряд новых переменных в базе данных (**_pscore**, **_weight** и др.), которые используются программой для проверки баланса и включаются в последующий анализ.

Далее мы проверим, насколько сбалансированным оказалось распределение конфаундеров между основной и контрольной группами после подбора пар 1 : 1:

```
psstest gender age length diam, treated(group) both graph
```

На основании представленных программой таблицы и графика (рис. 13) мы можем оценить степень уменьшения смещения средних значений всех 4 ковариат после проведения процедуры подбора пар (графа «%reduct |bias|»): 62,7 % для ковариаты *gender*; 16,4 % для ковариаты *age*; 45,2 % для ковариаты *length* и 69,6 % для ковариаты *diam*. Графически данная корректировка выглядит как приближение стандартизированной доли смещения для каждой ковариаты («Standardized % bias across covariates») к нулевому значению смещения после процедуры подбора пар или «взвешивания». Таким образом, чем более сбалансированы ковариаты, тем ближе к нулевому значению располагаются соответствующие точки на графике после использования метода подбора пар или «взвешивания».

Результат такого способа подбора пар в нашем случае оказался неудовлетворительным – стандартизированное отклонение («%bias») после подбора пар для некоторых ковариат оказалось значительно превышающим 10 % (пороговое значение, определяющее адекватность балансировки) [11].

5.2. Метод подбора пар с помощью поиска «ближайшего соседа», соответствующего заданному отрезку PS

С помощью данного метода мы повысим качество подбора пар, то есть сделаем так, чтобы наблюдения в парах были максимально похожи друг на друга по имеющимся конфаундерам.

Для этого потребуется задать значение отрезка (**caliper**), в рамках которого наблюдения основной группы и группы сравнения будут считаться эквивалентными. В нашем примере по-прежнему выбрано соотношение каждого наблюдения основной группы с 1-м наблюдением группы сравнения.

Значение данного отрезка определяется логитом PS. Для расчета логита PS потребуется последовательно ввести следующие две команды:

```
gen logitpscore = ln(mypscore / (1-mypscore))
```

```
sum logitpscore
```

Первая команда (**gen**) рассчитывает значения

Variable	Unmatched Matched	Mean		%bias	%reduct bias
		Treated	Control		
gender	U	.85135	.92373	-22.9	62.7
	M	.85135	.87838	-8.6	
age	U	69.081	65.364	52.4	16.4
	M	69.081	65.973	43.8	
length	U	30.419	24.72	34.3	45.2
	M	30.419	27.297	18.8	
diam	U	22.716	28.636	-101.2	69.6
	M	22.716	24.514	-30.7	

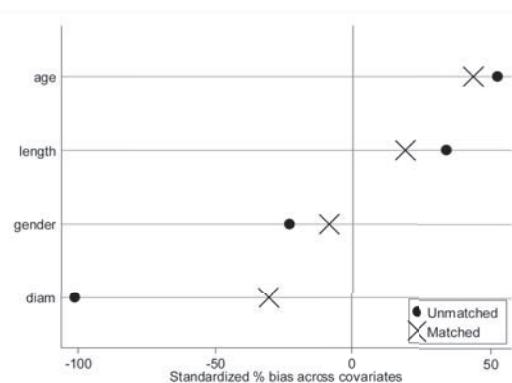


Рис. 13. Результаты проверки смещения конфаундеров в основной группе и группе контроля до (U) и после (M) подбора пар с использованием метода поиска «ближайшего соседа» 1 : 1

логита PS и создает соответствующую переменную, а вторая (**sum**) показывает короткую таблицу (не приводится), в которой программа укажет стандартное отклонение («Std. Dev.») логита PS (в нашем примере – 1,393). Установлено, что значение отрезка, равное 0,2 от стандартного отклонения логита PS, является достаточным для обеспечения довольно близкого соотношения наблюдений основной группы и группы сравнения [7, 11].

Соответственно 0,2 стандартных отклонения логита PS будут равны $0,2 \times 1,393 = 0,28$, и данную цифру можно вставить в команду как величину отрезка. Но в нашем случае значение 0,28 не позволяет достичь удовлетворительного баланса ковариат (читатель может убедиться в этом самостоятельно), поэтому будем уменьшать его величину и опытным путем дойдем до значения 0,04, которое позволяет адекватно сбалансировать ковариаты:

```
psmatch2 group, pscore(mypscore)
common caliper(0.04) neighbor (1)
noreplacement
```

На этот раз результаты подбора пар оказались удовлетворительными – стандартизованное отклонение ковариат составило не более 10 % (рис. 14).

5.3. Метод «взвешивания» Кернела

Несмотря на то, что адекватная балансировка ковариат была достигнута, продемонстрируем еще один способ балансировки, при котором каждому наблю-

дению контрольной группы придается определенный «вес» в зависимости от того, насколько оно близко к наблюдению основной группы по значению PS:

```
psmatch2 group, kernel pscore(mypscore)
```

На этот раз стандартизованное отклонение превысило значение 10 % для ковариаты age (рис. 15). Тем не менее, хотя «взвешивание» Кернела и не дало удовлетворительных результатов балансировки ковариат, мы используем его для дальнейшего анализа с целью тренировки.

6. Расчет меры эффекта действия изучаемого фактора

Этот этап является завершающим этапом обработки данных и позволяет непосредственно оценить меру эффекта «лечения» (фактора) с коррекцией на действие конфаундеров.

Одной из главных используемых мер эффекта действия фактора в случае бинарного исхода является отношение шансов (odds ratio – OR) – дробь, в числителе которой находятся шансы развития события в основной группе, а в знаменателе – шансы развития события в группе сравнения (или наоборот, в зависимости от того, какая из двух групп выступает в качестве контрольной).

В нашем примере мы сравниваем шансы развития осложнений при использовании одного метода оперативного вмешательства (закодирован цифрой «1») по сравнению с другим (цифра «0»).

Variable	Unmatched Matched	Mean		%bias	%reduct bias
		Treated	Control		
gender	U	.85135	.92373	-22.9	
	M	.86957	.8913	-6.9	70.0
age	U	69.081	65.364	52.4	
	M	67.348	67.022	4.6	91.2
length	U	30.419	24.72	34.3	
	M	29.304	29.696	-2.4	93.1
diam	U	22.716	28.636	-101.2	
	M	24.043	23.457	10.0	90.1

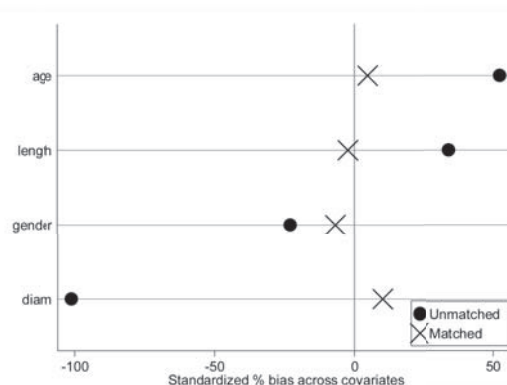


Рис. 14. Результаты проверки смещения конфаундеров в основной группе и группе контроля до (U) и после (M) подбора пар с использованием метода поиска «ближайшего соседа» 1 : 1 с заданным значением отрезка 0,04

Variable	Unmatched Matched	Mean		%bias	%reduct bias
		Treated	Control		
gender	U	.85135	.92373	-22.9	
	M	.85135	.86897	-5.6	75.7
age	U	69.081	65.364	52.4	
	M	69.081	68.084	14.1	73.2
length	U	30.419	24.72	34.3	
	M	30.419	30.86	-2.7	92.3
diam	U	22.716	28.636	-101.2	
	M	22.716	22.456	4.4	95.6

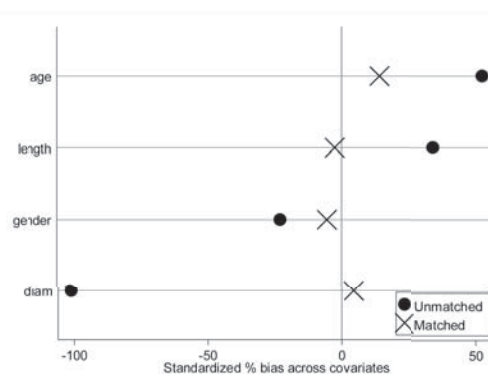


Рис. 15. Результаты проверки смещения конфаундеров в основной группе и группе контроля до (U) и после (M) «взвешивания» Кернела

Рассчитать значение OR, скорректированное на действие конфаундеров, мы можем с помощью нескольких нижеприведенных способов.

6.1. Корректировка результатов логистической регрессии с использованием PSM

6.1.1. Включение в анализ значений PS в качестве единственной ковариаты

Логистическая регрессия позволяет рассчитать OR для переменной «лечение» (*group*), при этом корректировка на действие конфаундеров производится путем включения в модель вместо всех ковариат только значения PS, которое, по сути, и заключает в себе интегрирующую все ковариаты переменную. Заметим, что при данном способе использования PS балансировка ковариат не проводится.

Для запуска расчетов введем команду:

```
logistic complicat group mypscore
```

Результаты представлены на рис. 16: отношение шансов («Odds Ratio») и его 95% доверительный интервал («[95% Conf. Interval]») представлены в той строке таблицы, в которой указана переменная «лечение» (*group*).

Так как доверительный интервал для OR не включает в себя единицу, мы можем сделать вывод о том, что полученные результаты статистически значимы: OR = 3,2 (95 % ДИ: 1,3–8,3).

Соответственно на основании результатов анализа мы можем заключить, что при выполнении оперативного вмешательства, закодированного в переменной *group* цифрой «1» (основная группа), шансы развития осложнений в 3,2 раза выше, чем при выполнении операции, закодированной цифрой «0» (контрольная группа).

Для повышения точности оценки мы можем

ограничить выборку наблюдений для логистической регрессии зоной «перекреста»:

```
logistic complicat group mypscore if comsup==1
```

Результаты расчетов представлены на рис. 17: OR = 3,2 (95% ДИ: 1,2–8,2). Заметим, что в этом случае количество включенных в анализ наблюдений уменьшилось до 175, и рассчитанное отношение шансов практически не изменилось по сравнению с полученным при использовании всей выборки. В данном случае целесообразно искать «золотую середину» между повышением точности расчета отношения шансов, что обеспечивается зоной «перекреста», и точностью логистической регрессии, для которой требуется включение в анализ как можно большего числа наблюдений.

6.1.2. Логистическая регрессия на выборке, полученной при подборе пар 1 : 1

Этот вариант анализа можно использовать только непосредственно после ввода представленной ранее команды:

```
psmatch2 group, pscore (myscore) common caliper (0.04) neighbor (1) noreplacement
```

На основании созданной этой командой дополнительной переменной *_weight* мы можем сформировать выборку, в которую войдут только те наблюдения, которые образуют пары, имеющие достаточно близкие значения PS. Теперь логистическую регрессию мы проведем на сформированной таким способом выборке, включив в регрессионную модель только переменные исхода (*complicat*) и «лечения» (*group*):

```
logistic complicat group if _weight==1
```

Результаты анализа представлены на рис. 18.

В данном случае 95 % доверительный интервал

Logistic regression							Number of obs	=	192
Log likelihood = -81.623539							LR chi2(2)	=	6.51
							Prob > chi2	=	0.0385
							Pseudo R2	=	0.0384
complicat	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]				
group	3.229793	1.553416	2.44	0.015	1.258279 8.290343				
myscore	.435152	.4303394	-0.84	0.400	.0626402 3.022934				
_cons	.1552242	.060381	-4.79	0.000	.0724184 .3327133				

Рис. 16. Результаты логистической регрессии при включении в модель значений PS в качестве единственной ковариаты

Logistic regression							Number of obs	=	175
Log likelihood = -75.455946							LR chi2(2)	=	6.25
							Prob > chi2	=	0.0440
							Pseudo R2	=	0.0397
complicat	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]				
group	3.195602	1.53859	2.41	0.016	1.243723 8.210727				
myscore	.4120066	.4307887	-0.85	0.396	.0530761 3.198228				
_cons	.1614578	.0722798	-4.07	0.000	.0671431 .3882546				

Рис. 17. Результаты логистической регрессии при включении в модель значений PS в качестве единственной ковариаты (выборка ограничена зоной «перекреста»)

```

Logistic regression
Log likelihood = -38.55238
Number of obs = 92
LR chi2(1) = 1.36
Prob > chi2 = 0.2428
Pseudo R2 = 0.0174
    
```

complicat	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
group	1.994594	1.200952	1.15	0.251	.6128346	6.491809
_cons	.1219513	.0577681	-4.44	0.000	.0481918	.3086023

Рис. 18. Результаты логистической регрессии с использованием выборки, сформированной методом подбора пар

для значения OR включает в себя единицу, то есть результаты не являются статистически значимыми: OR = 2,0 (95 % ДИ: 0,6–6,5).

6.1.3. Логистическая регрессия с использованием «весовых» коэффициентов, полученных в результате проведения «взвешивания» Кернела

Напомним, что «взвешивание» Кернела в нашем случае не дало удовлетворительных результатов балансировки ковариат, но тем не менее приведем формулу расчета OR для данного случая:

```

logistic complicat group gender age
length diam [pweight=_weight]
    
```

6.2. Метод Мантеля – Хензеля с использованием блоков PS

Данный метод предназначен для расчета значения OR, «взвешенного» с учетом стратификации групп наблюдений, при условии, что внутри страт конфаундеры являются сбалансированными между основной и контрольной группами.

Вспомним, что еще на этапе расчета значений PS с помощью команды **pscore** были сформированы сбалансированные блоки значений PS (страты), на основании которых мы и используем метод Мантеля – Хензеля:

```

cc complicat group, by (myblock)
    
```

Программа выведет таблицу с результатами расчета OR и его 95% доверительного интервала для каждого из 5 блоков – страт (рис. 19). В нижней части таблицы в строке «M-H combined» приводится скорректированное значение отношения шансов: OR = 3,5 (95 % ДИ: 1,3–9,5).

Number of block	OR	[95% Conf. Interval]		M-H Weight
1	10.8	.433668	659.6636	.1428571
2	1.142857	.0929115	9.074313	1.076923
3	7.388889	.7860326	350.1129	.4
4	.	.8093936	.	0
5	.3333333	.0036497	39.14118	.6
Crude	2.62987	1.078185	6.615765	
M-H combined	3.490149	1.277332	9.536391	

Test of homogeneity (Tarone) chi2(4) = 6.54 Pr>chi2 = 0.1625

Test that combined OR = 1:

```

Mantel-Haenszel chi2(1) = 6.49
Pr>chi2 = 0.0109
    
```

Рис. 19. Результаты использования метода Мантеля – Хензеля

Итак, мы оценили шансы развития осложнений при использовании одного варианта оперативного вмешательства по сравнению с другим, применив для корректировки результатов различные сочетания PSM, логистической регрессии и других методов.

Несомненно, определенный интерес представляет сравнение полученных результатов с теми, которые мы получили бы при использовании «классического» метода, используемого в таких случаях – логистической регрессии при включении в модель всех ковариат-конфаундеров.

Сначала для сравнения рассчитаем нескорректированную меру эффекта, включив в модель только переменные исхода и «лечение»:

```

logistic complicat group
    
```

Согласно результатам расчетов, представленным на рис. 20, вид оперативного вмешательства, использованный у пациентов основной группы, повышает шансы появления местных осложнений в 2,7 раза (95 % ДИ: 1,3–6,0) по сравнению с вмешательством, использованных у пациентов группы контроля, что существенно ниже меры эффекта, полученной при использовании корректировки на действие конфаундеров.

Далее рассмотрим уже скорректированную модель, для чего включим в анализ все ковариаты:

```

logistic complicat group gender age
length diam
    
```

Согласно результатам расчетов, представленным рис. 21, OR = 3,3 (95 % ДИ: 1,3–8,6).

Согласно результатам расчетов (рис. 19) скорректированное OR = 3,3 (95 % ДИ: 1,3–8,6).

Результаты расчетов OR при использовании вышеречисленных методов представлены в таблице.

Результаты расчета отношения шансов с использованием различных методов

Метод	Отношение шансов	95% доверительный интервал	
		Нижняя граница	Верхняя граница
Логистическая регрессия без коррекции на действие конфаундеров	2,7	1,3	6,0
Логистическая регрессия с коррекцией на действие конфаундеров	3,3	1,3	8,6
Логистическая регрессия с включением в модель значений PS	3,2	1,3	8,3
Логистическая регрессия с включением в модель значений PS с учетом зоны «перекреста»	3,2	1,2	8,2
Логистическая регрессия на выборке, сформированной методом подбора пар 1 : 1	2,0	0,6	6,5
Логистическая регрессия на «взвешенной» выборке	4,4	1,4	13,7
Метод Мантеля – Хензеля с использованием блоков PS	3,5	1,3	9,6

Если рассматривать результаты корректировки с помощью PSM, то включение значений PS в логистическую регрессионную модель в качестве единственной ковариаты и метод Мантеля – Хензеля демонстрируют очень похожие результаты, и к тому же они вполне сопоставимы с результатами использованного «классического» способа коррекции – логистической регрессии.

Логистическая регрессия на выборке, полученной методом подбора пар 1 : 1, показала статистически незначимые результаты, что объясняется значительным уменьшением размеров выборки, и соответственно потерей статистической мощности, необходимой для получения корректных результатов логистической регрессии (при большом размере выборки данное явление будет нивелировано).

Логистическая регрессия на «взвешенной» выборке была проведена исключительно с тренировочной целью и показала завышенный результат, что связано с недостаточной сбалансированностью конфаундеров на этапе «взвешивания» наблюдений. Данный факт не исключает возможности использования данного способа коррекции исследователями, так как невозможно заранее предугадать, какой из методов PSM позволит наиболее адекватно сбалансировать конфаундеры между основной и контрольной группами.

В заключение следует сказать, что представленный в статье достаточно большой арсенал способов проведения псевдорандомизации с использованием PSM позволит исследователю выбрать подходящий и успешно провести анализ собственных исследовательских данных.

Список литературы

1. Гржибовский А. М., Иванов С. В. Когортные исследования в здравоохранении // Наука и Здравоохранение. 2015. № 3. С. 5–16.
2. Гржибовский А. М., Иванов С. В. Поперечные (одномоментные) исследования в здравоохранении // Наука и Здравоохранение. 2015. № 2. С. 5–18.
3. Унгуряну Т. Н., Гржибовский А. М. Программное

Logistic regression	Number of obs =	193
	LR chi2(1) =	6.62
	Prob > chi2 =	0.0101
	Pseudo R2 =	0.0382
Log likelihood = -83.378181		

complicat	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
group	2.740385	1.085606	2.54	0.011	1.260687 5.956839
_cons	.1238095	.0364023	-7.11	0.000	.0695801 .2203043

Рис. 20. Результаты использования логистической регрессии без корректировки на действие конфаундеров

Logistic regression	Number of obs =	192
	LR chi2(5) =	12.01
	Prob > chi2 =	0.0346
	Pseudo R2 =	0.0708
Log likelihood = -78.872591		

complicat	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
group	3.319144	1.61011	2.47	0.013	1.282656 8.588992
gender	2.034935	1.606562	0.90	0.368	.433047 9.562386
age	.9685809	.0284569	-1.09	0.277	.9143819 1.025992
length	1.020035	.0114883	1.76	0.078	.9977648 1.042802
diam	1.030241	.0366112	0.84	0.402	.9609263 1.104556
_cons	.1249654	.2952045	-0.88	0.379	.001219 12.81074

Рис. 21. Результаты использования логистической регрессии с корректировкой на действие конфаундеров

обеспечение для статистической обработки данных STATA: введение // Экология человека. 2014. № 1. С. 60–63.

4. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. М. : Медиа Сфера, 1998. 352 с.

5. Acock A. C. *Gentle Introduction to Stata*. USA, Texas : Stata Press, 2006. 289 p.

6. Austin P. C. An introduction to propensity score methods for reducing the effects of confounding in observational studies // *Multivariate Behavioral Research*. 2011. Vol. 46. P. 399–424.

7. Austin P. C. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies // *Pharmaceutical statistics*. 2011. Vol. 10 (2). P. 150–161.

8. Becker S.O., Ichino A. Estimation of average treatment effects based on propensity scores // *The Stata Journal*. 2002. Vol. 2 (4). P. 358–377.

9. Cepeda M. S., Boston R., Farrar J. T., Strom B. L. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders // *American Journal of Epidemiology*. 2003. Vol. 158 (3). P. 280–287.

10. D'Agostino R. B. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group // *Statistics in Medicine*. 1998. Vol. 17 (19). P. 2265–2281.

11. Garrido M. M., Kelley A. S., Paris J., Roza K., Meier D. E., Morrison R. S., Aldridge M. D. Methods for constructing and assessing propensity scores // *Health Services Research*. 2014. Vol. 49 (5). P. 1701–1720.

12. Guo Sh. Y., Mark W. Fraser. *Propensity Score Analysis: Statistical Methods and Applications*, 2nd ed. USA. SAGE Publications, 2015. 448 p.

13. Hernan M. A., Hernandez-Diaz S., Robins J. M. A structural approach to selection bias // *Epidemiology*. 2004. Vol. 15 (5). P. 615–625.

14. Patorno E., Grotta A., Bellocco R., Schneeweiss S. Propensity score methodology for confounding control in health care utilization databases // *Epidemiology Biostatistics and Public Health*. 2013. Vol. 10 (3). P. e8940.

15. Rosenbaum P. R., Rubin D. B. The central role of the propensity score in observational studies for causal effects // *Biometrika*. 1983. Vol. 70 (1). P. 41–55.

16. Sturmer T., Joshi M., Glynn R. J., Avorn J., Rothman K. J., Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods // *Journal of Clinical Epidemiology*. 2006. Vol. 59 (5). P. 437–447.

References

1. Grjibovski A. M., Ivanov S. V. Cohort studies in health sciences. *Nauka i Zdravoohranenie* [Science & Healthcare]. 2015, 3, pp. 5-16. [in Russian].

2. Grjibovski A. M., Ivanov S. V. Cross-sectional studies in health sciences. *Nauka I Zdravoohranenie* [Science & Healthcare]. 2015, 2, pp. 5-18. [in Russian]

3. Unguryanu T. N., Grjibovski A. M. Introduction to STATA - software for statistical analysis. *Ekologiya cheloveka* [Human Ecology]. 2014, 1, pp. 60-63. [in Russian]

4. Fletcher R., Fletcher C., Vagner E. *Klinicheskaya epidemiologiya. Osnovy dokazatel'noi meditsiny* [Clinical

epidemiology. Basics of the evidence-based medicine]. Moscow, 1998, 352 p.

5. Acock A. C. *Gentle Introduction to Stata*. USA, Texas, Stata Press, 2006, 289 p.

6. Austin P. C. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*. 2011, 46, pp. 399-424.

7. Austin P.C. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics*. 2011, 10 (2), pp. 150-161.

8. Becker S. O., Ichino A. Estimation of average treatment effects based on propensity scores. *The Stata Journal*. 2002, 2 (4), pp. 358-377.

9. Cepeda M. S., Boston R., Farrar J. T., Strom B. L. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American Journal of Epidemiology*. 2003, 158 (3), pp. 280-287.

10. D'Agostino R. B. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*. 1998, 17 (19), pp. 2265-2281.

11. Garrido M. M., Kelley A. S., Paris J., Roza K., Meier D. E., Morrison R. S., Aldridge M. D. Methods for constructing and assessing propensity scores. *Health Services Research*. 2014, 49 (5), pp. 1701-1720.

12. Guo Sh. Y., Mark W. Fraser. *Propensity Score Analysis: Statistical Methods and Applications*, 2nd ed. USA. SAGE Publications, 2015, 448 p.

13. Hernan M. A., Hernandez-Diaz S., Robins J. M. A structural approach to selection bias. *Epidemiology*. 2004, 15 (5), pp. 615-625.

14. Patorno E., Grotta A., Bellocco R., Schneeweiss S. Propensity score methodology for confounding control in health care utilization databases. *Epidemiology Biostatistics and Public Health*. 2013, 10 (3), p. e8940.

15. Rosenbaum P. R., Rubin D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983, 70 (1), pp. 41-55.

16. Sturmer T., Joshi M., Glynn R. J., Avorn J., Rothman K. J., Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *Journal of Clinical Epidemiology*. 2006, 59 (5), pp. 437-447.

Контактная информация:

Гржибовский Андрей Мечиславович — доктор медицины, старший советник Национального института общественного здравоохранения, г. Осло, Норвегия; руководитель отдела международных программ и инновационного развития ЦНИЛ СГМУ, г. Архангельск, Россия; профессор кафедры общественного здоровья и здравоохранения медицинского института Северо-Восточного федерального университета, г. Якутск, Россия; профессор, почетный доктор Международного казахско-турецкого университета, г. Туркестан, Казахстан; почетный профессор ГМУ г. Семей, Казахстан
Адрес: INFA, Nasjonalt folkehelseinstitutt, Postboks 4404 Nydalen, 0403 Oslo, Norway.

Тел.: +4745268913 (Норвегия), +79214717053 (Россия), +77471262965 (Казахстан)

E-mail: Andrej.Grjibovski@gmail.com