



Возможности использования направленных ациклических графов при планировании и интерпретации результатов биомедицинских исследований

Е.А. Кригер¹, В.А. Постоев¹, А.В. Кудрявцев¹, Т.Н. Унгурияну¹, А.М. Гржибовский^{1,2,3}

¹ Северный государственный медицинский университет, Архангельск, Россия;

² Северо-Восточный федеральный университет им. М.К. Аммосова, Якутск, Россия;

³ Северный (Арктический) федеральный университет им. М.В. Ломоносова, Архангельск, Россия

АННОТАЦИЯ

В статье представлен алгоритм построения и применения направленных ациклических графов (Directed Acyclic Graphs, DAGs) при планировании эпидемиологических и биомедицинских исследований. DAGs представляют собой графический инструмент для моделирования сложных связей между переменными, что особенно актуально в биомедицинской науке, где корректная оценка причинно-следственных связей требует учёта потенциальных вмешивающихся факторов. Подчёркивается значимость DAGs для концептуализации научной гипотезы и понимания структуры связей между факторами на основе анализа литературы и результатов ранее проведённых исследований. Применение DAGs способствует повышению качества как планирования исследования, так и анализа данных, обеспечивая более обоснованный подход к отбору переменных для включения в математические модели. DAGs позволяют определить минимальный и достаточный набор факторов для коррекции с учётом роли переменных (конфаундеров, медиаторов, коллайдеров) относительно воздействия (вероятного фактора риска) и исхода (заболевания или состояния), тем самым снижая риск аналитических ошибок. В статье акцентируется внимание на практическом применении DAGs с помощью доступного программного обеспечения, а также представлены конкретные примеры использования графов в биомедицинских исследованиях. В заключение предложены рекомендации по интеграции DAGs в практику биомедицинских исследований, что может способствовать более широкому внедрению современных методов многомерного статистического анализа, улучшению интерпретируемости и повышению воспроизводимости научных результатов.

Ключевые слова: направленные ациклические графы; причинно-следственная связь; наблюдательные исследования; эпидемиология; конфаундер; медиатор; коллайдер.

Как цитировать:

Кригер Е.А., Постоев В.А., Кудрявцев А.В., Унгурияну Т.Н., Гржибовский А.М. Возможности использования направленных ациклических графов при планировании и интерпретации результатов биомедицинских исследований // Экология человека. 2025. Т. 32, № 5. С. 300–314.

DOI: 10.17816/humeco683466 EDN: MXPLRC

Potential Applications of Directed Acyclic Graphs in the Design and Interpretation of Biomedical Research

Ekaterina A. Krieger¹, Vitaly A. Postoev¹, Alexander V. Kudryavtsev¹,
Tatiana N. Unguryanu¹, Andrej M Grjibovski^{1,2,3}

¹ Northern State Medical University, Arkhangelsk, Russia;

² North-Eastern Federal University named after M.K. Ammosov, Yakutsk, Russia;

³ Northern (Arctic) Federal University named after M.V. Lomonosov, Arkhangelsk, Russia

ABSTRACT

This article presents an algorithm for constructing and applying directed acyclic graphs (DAGs) in the planning of epidemiological and biomedical studies. DAGs are graphical tools for modeling complex relationships between variables, which is particularly relevant in biomedical science, where accurate assessment of causal relationships requires accounting for potential confounding factors. The importance of DAGs is emphasized for conceptualizing scientific hypotheses and understanding the structure of relationships between factors based on scientific data review and findings from previous studies. The use of DAGs enhances the quality of both study design and data analysis, providing a more grounded approach to selecting variables for inclusion in statistical models. DAGs make it possible to determine the minimal and sufficient set of factors for adjustment, with consideration of the roles of variables (confounders, mediators, colliders) in relation to the exposure (a probable risk factor) and the outcome (a disease or condition), thus reducing the likelihood of analytical errors. The article highlights the practical application of DAGs using available software and provides specific examples of their use in biomedical research. Finally, recommendations are offered for integrating DAGs into biomedical research practice, which may contribute to the broader adoption of modern multivariate statistical methods, improved interpretability, and enhanced reproducibility of scientific findings.

Keywords: directed acyclic graphs; causal inference; observational studies; epidemiology; confounder; mediator; collider.

To cite this article:

Krieger EA, Postoev VA, Kudryavtsev AV, Unguryanu TN, Grjibovski AM. Potential applications of directed acyclic graphs in the design and interpretation of biomedical research exposure risk and the assessing methodology for its increase. *Ekologiya cheloveka (Human Ecology)*. 2025;32(5):300–314. DOI: 10.17816/humeco683466 EDN: MXPLRC

DOI: <https://doi.org/10.17816/humeco683466>

EDN: MXPLRC

在生物医学研究的设计与结果解释中应用有向无环图的可能性

Ekaterina A. Krieger¹, Vitaly A. Postoev¹, Alexander V. Kudryavtsev¹,
Tatiana N. Unguryanu¹, Andrej M. Grjibovski^{1,2,3}

¹ Northern State Medical University, Arkhangelsk, Russia;

² North-Eastern Federal University named after M.K. Ammosov, Yakutsk, Russia;

³ Northern (Arctic) Federal University named after M.V. Lomonosov, Arkhangelsk, Russia

摘要

本文介绍了一种在流行病学和生物医学研究中构建与应用有向无环图 (Directed Acyclic Graphs, DAGs) 的算法。DAGs 是一种用于建模变量之间复杂关系的图示工具，尤其适用于生物医学领域，在该领域中，对因果关系的准确评估依赖于对潜在混杂因素的充分考虑。DAGs 在科学假设的构建与变量关系结构的理解中具有重要意义，其依据包括文献分析与既有研究成果。应用 DAGs 可提高研究设计与数据分析的质量，为变量选择提供理论依据，从而优化数学模型的构建过程。DAGs 可用于在因果结构中明确变量（如混杂变量、中介变量、碰撞变量）相对于暴露（潜在风险因素）与结局（疾病或状态）的作用，从而识别进行校正所需的最小充分变量集，并降低分析误差风险。文章强调了利用现有软件工具实践应用 DAGs 的重要性，并展示了其在生物医学研究中具体使用的实例。最后，提出将 DAGs 纳入生物医学研究流程的建议，以促进现代多变量统计分析方法的推广，提高研究结果的可解释性和重复性。

关键词：有向无环图；因果关系；观察性研究；流行病学；混杂变量；中介变量；碰撞变量。

引用本文：

Krieger EA, Postoev VA, Kudryavtsev AV, Unguryanu TN, Grjibovski AM. 在生物医学研究的设计与结果解释中应用有向无环图的可能性. *Ekologiya cheloveka (Human Ecology)*. 2025;32(5):300–314. DOI: 10.17816/humeco683466 EDN: MXPLRC

收到: 09.06.2025

接受: 22.06.2025

发布日期: 10.07.2025

ОБОСНОВАНИЕ

Оценка причинно-следственной связи между воздействием (фактор риска) и исходом (заболевание или состояние) — одна из ключевых задач эпидемиологических исследований. При этом задачей исследователя является оценка независимого влияния воздействия на исход с исключением влияния на изучаемую связь других факторов, которые носят название конфаундеров, или смешивающих факторов. Наиболее надёжным методом определения причинно-следственной связи является рандомизированное контролируемое испытание, поскольку процедура рандомизации теоретически обеспечивает равномерное распределение участников с различными характеристиками между группами, что устраняет влияние потенциальных конфаундеров на изучаемую связь. В то же время рандомизация возможна только при экспериментальных исследованиях, а их использование, в свою очередь, ограничено высокой стоимостью, трудоёмкостью, невозможностью изучения влияния вредных воздействий. В связи с этим исследователи чаще прибегают к анализу связей на основе данных наблюдательных (обсервационных) исследований, таких как когортные или исследования по типу «случай–контроль». Поперечный дизайн исследования, несмотря на ограничения, может быть использован для оценки ассоциаций между воздействием и исходом и формирования гипотез, которые в дальнейшем могут быть проверены с использованием более подходящих для установления причинно-следственной связи дизайнов, описанных выше [1]. При изучении причинно-следственной связи в наблюдательных исследованиях требуется глубокое понимание роли различных факторов в отношении воздействия и исхода для минимизации влияния этих факторов на изучаемую связь и более корректной её оценки и интерпретации. Например, проживание в сельской местности (изучаемое воздействие) при первоначальном грубом или нескорректированном анализе может быть связано с риском смерти от острых состояний (изучаемый исход), связанных с нарушением кровообращения, но при учёте вмешивающихся факторов (более высокий средний возраст жителей села, большая вероятность несвоевременного оказания медицинской помощи) сила связи уменьшается до незначимых уровней.

Направленные ациклические графы (Directed Acyclic Graphs — DAGs) — это графический инструмент, который используется в наблюдательных эпидемиологических исследованиях для анализа связи (причинно-следственной) между изучаемым воздействием и исходом с учётом возможного влияния на силу и направленность этой связи других потенциально вмешивающихся факторов [2–4].

DAGs представляют собой простой способ визуализации связи между воздействием и исходом на основании анализа имеющихся литературных данных, а также знаний, гипотез и допущений исследователя [2–4]. Использование DAGs позволяет структурировать информацию,

полученную по результатам ранее проведённых исследований, чётко сформулировать исследовательский вопрос, наглядно представить связи между различными факторами и определить минимальный набор потенциально вмешивающихся факторов, которые необходимо учитывать для корректной оценки влияния воздействия на исход. Всё это повышает воспроизводимость исследования, делая его более понятным для широкого круга исследователей, улучшает интерпретируемость и сопоставимость его результатов с результатами других исследований.

Несмотря на преимущества DAGs для визуализации, корректной оценки и интерпретации изучаемых связей, их использование в российских биомедицинских исследованиях остаётся ограниченным. Вероятно, это связано с недостаточной осведомлённостью исследователей о возможностях данного метода и отсутствием рекомендаций на русском языке по построению, интерпретации и представлению DAGs в публикациях. В настоящей статье представлены пошаговые инструкции по использованию DAGs в биомедицинских исследованиях с использованием как русскоязычной, так и англоязычной терминологии, которая поможет отечественным авторам работать с англоязычным программным обеспечением и лучше понимать иностранные научные публикации.

ТЕРМИНОЛОГИЯ DAGs

DAGs — это непараметрическое графическое представление набора факторов (переменных) в заданном контексте исследования. В терминологии DAGs изучаемые факторы (переменные) называются узлами (nodes), которые соединяются между собой дугами/стрелками (arcs/arrows; рис. 1). Стрелка между двумя узлами обозначает наличие и направление предполагаемой (причинно-следственной) связи, но не указывает её знак (положительный или отрицательный), величину (сильная или слабая), характер (линейная или нелинейная) или точное математическое выражение, что делает DAGs непараметрической моделью.

Узлы, от которых исходят стрелки, называются предками (ancestors), а узлы, к которым направлены стрелки, — потомками (descendants). Для улучшения читаемости и интерпретируемости графа рекомендуется соблюдать направленность стрелок в одном направлении (как правило, слева направо), а также избегать перекрещивания стрелок. Важно не перегружать граф избыточными узлами, не влияющими на анализируемую (причинно-следственную) связь между воздействием и исходом.

Говоря о наличии связи между воздействием и исходом, мы предполагаем, что изменение воздействия приведёт к изменению вероятности исхода — вероятностное обоснование (probabilistic reasoning). В контексте контрфактического обоснования (counterfactual reasoning) мы предполагаем, что если воздействие изменится, то изменится и исход.

Последовательность соединённых стрелками узлов, ведущая от воздействия к исходу, обозначается термином «путь» (path). Ни один из факторов (переменных) не может быть причиной самого себя, то есть не существует пути, который проходит дважды через один узел. Поэтому в DAGs нет замкнутых циклов, что отражает прямой порядок причинности: причина → следствие [5]. Разные пути могут проходить через один узел, а узел, в котором сходятся несколько стрелок, называется суперузлом (super node) [6].

Пути бывают открытыми и закрытыми. Путь между двумя переменными считается открытым, если между ними существует связь: знание значения одной переменной позволяет оценить вероятность другой. Путь считается закрытым, если значимой связи нет. Пути, которые

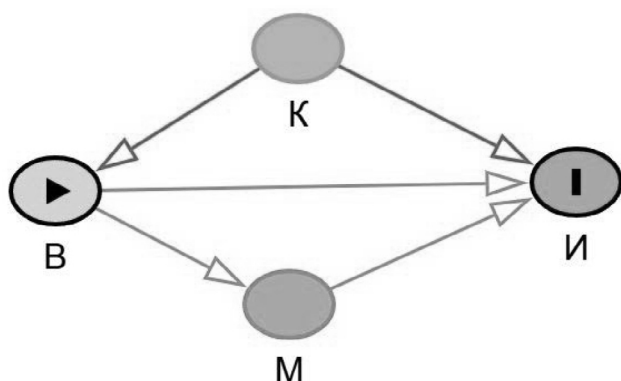


Рис. 1. Пример построения DAG.

Fig. 1. Example of DAG construction.

отображают реальное влияние одной переменной на другую, называются причинно-следственными. Пути, которые создают ложные связи между переменными, называются ошибочными (смещёнными) путями (biasing path). На рис. 1 причинно-следственные пути обозначены зелёными стрелками, а ошибочные пути — красными. Для точной оценки изучаемой связи важно закрыть ошибочные пути с помощью учёта и коррекции влиянием вмешивающихся факторов, связанных с воздействием и исходом посредством этих ошибочных путей.

ПОСТРОЕНИЕ DAGs

Построение DAGs возможно с использованием различных программных инструментов, включая инструмент DAGitty, пакеты ggdag и dagitty для R, библиотек networkX и rgmpy для Python [7]. В данной статье представлен пошаговый алгоритм построения DAGs с использованием программного обеспечения DAGitty. Программа DAGitty доступна бесплатно как в онлайн-версии через браузер на сайте www.dagitty.net, так и в виде приложения, которое можно установить на персональный компьютер. DAGitty разработана и поддерживается Институтом вычислительных и информационных наук Радбаудского университета (The Institute for Computing and Information Sciences at Radboud University) и кафедрой биомедицинских наук университетской клиники Радбауд (The Medical BioSciences department at Radboudumc) в Неймергене,

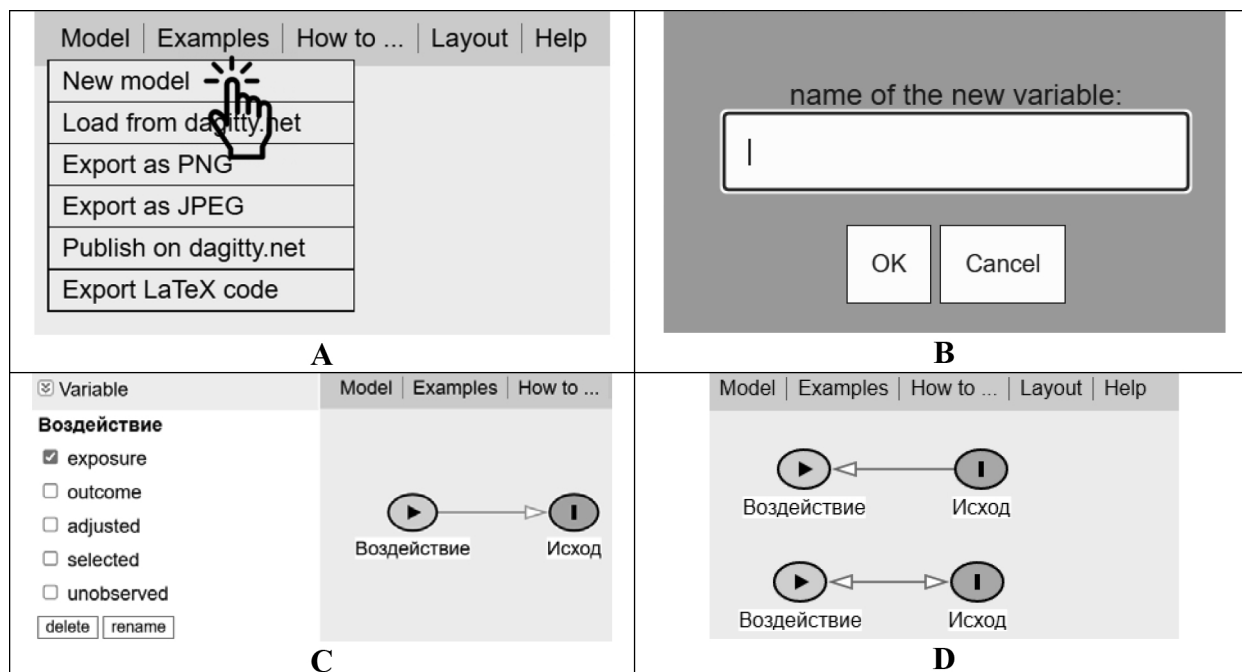


Рис. 2. Начало (инициация) построения DAGs, определение воздействия и исхода.

Перевод интерфейса программы: А — New model (новая модель), Load from dagitty.net (загрузить из dagitty.net), Export as PNG (экспортировать/скачать в формате PNG), Export as JPEG (экспортировать/скачать в формате JPEG), Publish on dagitty.net (публиковать в dagitty.net), Export LaTeX code (экспортировать/скачать LaTeX код); В — name of the new variable (название новой переменной), OK (да/принять), Cancel (отмена); С — Variable (Переменная), exposure (воздействие), outcome (исход), adjusted (скорректировано), selected (выбрано), unobserved (не наблюдается/нет данных).

Fig. 2. Initiation of DAG construction: defining the exposure and the outcome.

Нидерланды [7]. Язык программного обеспечения — английский.

Построение DAGs рекомендуется на этапе планирования исследования, до начала сбора данных, для определения критериев включения и исключения участников при их отборе, а также перечня потенциально вмешивающихся факторов (переменных), информацию о которых предстоит собрать. При невозможности построения DAGs на этапе планирования, например, когда исследователь работает с ранее собранными данными (вторичный анализ данных), построение DAGs следует выполнить перед проведением статистического анализа.

Для построения DAGs на официальном сайте www.dagitty.net в меню браузера (в верхней части экрана) нужно выбрать вкладку «Model» (Модель), а затем выбрать из выпадающего списка «New model» (Новая модель; рис. 2А), после чего появится специальная область на экране для построения графа. При клике в любой части этой области откроется окно, в которое нужно ввести название фактора (переменной; рис. 2В). Такую процедуру необходимо будет повторить для всех факторов (переменных), которые должны быть включены в анализ. При необходимости название переменных можно изменить. Для этого нужно кликнуть на переменную, которую надо переименовать, и выбрать «rename» (Переименовать) в меню «Variable» (Переменная) в левом верхнем углу экрана (рис. 2С).

Построение DAGs начинается с определения воздействия и исхода. Для этого нужно выбрать переменную, затем в меню «Variable» (Переменная) установить галочку рядом с «Exposure» (Воздействие) или «Outcome» (Исход), в зависимости от того, какую роль будет выполнять выбранная переменная в вашем анализе (рис. 2С). После определения переменных как воздействия или исхода они будут обозначаться специальными значками (зелёный для воздействия и синий для исхода), позволяющими отличать эти ключевые переменные от других переменных в графе.

Каждый граф должен включать одно воздействие и один исход. Если целью является оценка влияния нескольких воздействий на один исход, для каждого воздействия строится отдельный граф, поскольку на связь между каждым конкретным воздействием и исходом влияют уникальные факторы, которые важно выявить и учесть при анализе. Связи, представляющие интерес для исследователя, должны быть чётко отражены в цели исследования.

После того как воздействие и исход определены, необходимо обозначить связь между ними стрелкой, направленной от воздействия к исходу. Чтобы создать стрелку, нужно кликнуть на переменную «воздействие», а затем на переменную «исход». Если действие повторить, стрелка исчезнет. Зелёная стрелка указывает на причинный путь (то есть причинно-следственную связь), а красная — на ошибочную связь. Цвет стрелки определяется автоматически программой и отражает корректность

направленной связи. Если стрелка направлена от исхода к воздействию или становится двунаправленной, она окрасится в красный, указывая на ошибку (рис. 2Д). Это связано с тем, что DAGs должны быть ациклическими, замкнутые циклы не допускаются.

Далее необходимо добавить в граф факторы (переменные), потенциально связанные с воздействием и/или исходом, и указать стрелками направленность связи. В граф следует включить все переменные, которые по данным литературы или исходя из гипотезы и допущений исследователя важны для оценки связи. Даже при отсутствии данных о некоторых факторах их необходимо учитывать, если теоретически они могут влиять на изучаемую связь. При построении DAGs проверяется наличие связи между любыми двумя переменными. Отсутствие стрелки между переменными означает, что между ними нет связи (ассоциации).

Минимальный необходимый и достаточный для корректной оценки изучаемой (причинно-следственной) связи набор переменных определяется при построении DAGs визуально или алгоритмически. Точность полученной оценки зависит от того, насколько корректно DAGs отражает фактические механизмы связей между переменными, что требует критического анализа. В зависимости от цели исследования и заданных условий программа DAGitty может предложить несколько допустимых наборов переменных для коррекции. Для каждого из рекомендованных наборов можно поочерёдно выбирать переменные и устанавливать напротив них галочку «Adjusted» (Скорректировано) в меню «Variable» (Переменная), чтобы обозначить коррекцию. Это позволяет визуально оценить, как изменяются причинные пути в графе после коррекции. При правильной коррекции красные стрелки, обозначающие ошибочный путь, станут чёрными (ошибочный путь закрыт). Следовательно, итоговая модель не будет содержать красных стрелок.

Полученный с помощью DAGitty граф можно скачать в формате рисунка, выбрав в меню DAGitty «Model» (Модель), а затем выбрать из выпадающего списка «Export as PNG/JPEG» (Экспортировать в формате PNG/JPEG).

При построении DAGs автоматически формируется код модели (Model code), текстовое описание структуры графа, которое можно скопировать, сохранить, переслать коллегам. Такой формат упрощает воспроизведение и редактирование графа: не потребуется строить его заново — достаточно вставить сохранённый код в окно «Model code» (Код модели) в правой части рабочей области DAGitty и нажать «Update DAG» (Обновить DAG).

РОЛИ ПЕРЕМЕННЫХ ПО ОТНОШЕНИЮ К ВОЗДЕЙСТВИЮ И ИСХОДУ

Особое значение отводится определению роли факторов (переменных) по отношению к воздействию и исходу

для выявления факторов, учитывать которые необходимо, и факторов, коррекция на которые может привести к искажению изучаемой связи. На рис. 3 показаны основные компоненты DAGs и роли факторов (переменных), определяемые в контексте изучения (причинно-следственной) связи между воздействием и исходом. В терминологии DAGs факторы, влияющие на связь между воздействием и исходом, классифицируются как конфаундеры (confounders; рис. 3A), медиаторы/модификаторы эффекта (mediators/effect modifiers; рис. 3B) и коллайдеры (colliders; рис. 3C).

DAGs позволяют выделить конфаундеры, медиаторы (модификаторы эффекта) и коллайдеры, помогая понять, какие переменные следует учитывать при анализе данных.

Конфаундер — это переменная, которая связана как с воздействием, так и с исходом, при этом не являясь промежуточным звеном в причинно-следственной связи между ними, создавая ложную или искажая истинную связь между изучаемыми переменными [8]. Иными словами, между группой, подверженной изучаемому воздействию, и группой, не подверженной этому воздействию, существуют различия в частоте третьего фактора (конфаундера), который влияет на вероятность возникновения исхода. Например, изучение порядкового номера беременности как фактора риска возникновения синдрома Дауна у новорождённого может привести к выявлению наличия ложной причинно-следственной связи, если не учесть возраст матери как потенциальный конфаундер. Логично предположить, что с увеличением паритета возраст женщины увеличивается, то есть средний возраст женщин с большим количеством беременностей будет выше, чем у первородящих.

В терминологии DAGs путь конфаундинга или обратный путь (backdoor path) — это путь от воздействия к исходу, проходящий через конфаундер. Если влияние конфаундера не учтено (не проведена коррекция), этот путь является открытым и создаёт ложную связь между изучаемыми переменными. Коррекция на конфаундер делает данный путь закрытым.

Учесть влияние конфаундера можно как на этапе планирования исследования (ограничение, рандомизация), так и на этапе статистической обработки с использованием

таких методов, как стратификация/стандартизация и многомерный регрессионный анализ [9].

Ограничение подразумевает включение в исследование или проведение анализа с использованием данных только тех участников, которые соответствуют определённым характеристикам (критериям включения). Например, только некурящих (чтобы устранить эффект курения).

Стратификация — это метод, при котором выборка делится на подгруппы (страты) на основе значения конфаундера (например, возраст, пол или уровень дохода). Эффект воздействия оценивается в каждой из этих подгрупп отдельно.

Стандартизация используется для коррекции влияния конфаундера на оценку эффекта воздействия, при этом учитывается распределение конфаундера в стандартной или эталонной популяции. Целью проведения стандартизации является получение показателей, пригодных для сравнения в двух совокупностях, неоднородных по своему составу, то есть отличающихся по какому-либо количественному или категориальному признаку. Чаще всего стандартизацию используют для устранения влияния различий в возрастной структуре сравниваемых популяций, например, для сравнения частоты и смертности от возраст-ассоциированных заболеваний, таких как злокачественные новообразования, болезни системы кровообращения.

Чтобы провести коррекцию на конфаундеры с помощью многомерной регрессии, необходимо определить перечень переменных, минимально необходимый и достаточный для получения наиболее точной оценки при помощи DAGs.

Альтернативным DAGs подходом выявления конфаундеров, который часто используется исследователями, является подход, основанный на данных (data-driven approach). При использовании подхода, основанного на данных, переменные для включения в многомерную регрессионную модель выбираются на основе их статистической значимости в одномерных моделях, которые оценивают связь между воздействием и исходом [10, 11]. Переменные не включаются в многомерную модель и не считаются конфаундерами, если они не демонстрируют значимой связи в одномерных моделях. При этом в качестве критического уровня значимости в одномерной модели может использоваться значение альфа-ошибки

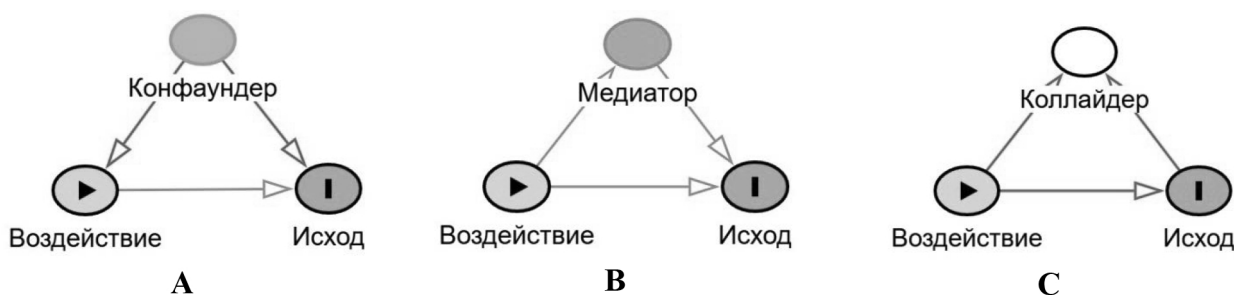


Рис. 3. Роли факторов (переменных) в контексте изучения причинно-следственной связи между воздействием и исходом.

Fig. 3. Roles of variables in the context of causal inference between exposure and outcome.

выше привычного 0,05, например, равное 0,1–0,2 [10]. Конфаундеры можно выявить при помощи оценки изменения величины эффекта воздействия на исход (the change-in-estimate procedure) при добавлении в модель потенциального конфаундера. Если величина эффекта изменяется на 10–20% или более после добавления оцениваемого фактора, добавленная переменная считается конфаундером [10, 12]. Хотя подход, основанный на данных, предоставляет ценную информацию о связи между переменными, он не позволяет с точностью определить, какую роль играет добавленная в модель переменная — конфаундер, медиатор или коллайдер. Наличие любого из этих факторов может повлиять на изучаемую связь, а их включение в регрессионную модель может повлиять на оценку эффекта. Следовательно, выбор переменных с использованием данного подхода может привести к ошибочному исключению или включению переменных в модель. В отличие от подхода, основанного на данных, использование DAGs позволяет определить роли ковариат (переменных) в регрессионной модели, включая идентификацию конфаундеров, медиаторов и коллайдеров [13]. В некоторых статьях используются гибридные подходы к выбору переменных, сочетающие DAG с традиционными статистическими методами.

Рассмотрим влияние конфаундера на конкретном примере. Анализируя данные электронных медицинских карт 64 000 детей, исследователи выявили связь между частотой приёма антибиотиков в течение первых двух лет жизни и развитием ожирения в более позднем возрасте, причём этот эффект был наиболее выражен для антибиотиков широкого спектра действия [14]. Позднее авторы другого исследования показали, что описанная ранее связь между приёмом антибиотиков и ожирением объясняется частотой инфекций (конфаундер), которые являются показанием к применению антибиотиков и могут влиять на риск развития ожирения [15]. Для учёта влияния конфаундера авторы разделили когорту детей (N=260 556) на страты в зависимости от частоты инфекционных заболеваний и применения антибиотиков в первые 12 мес. жизни: 1) дети, не болевшие инфекциями; 2) дети, перенёсшие инфекции, но не получавшие антибиотиков; 3) дети, перенёсшие инфекции и получавшие антибиотики. Инфекции, перенесённые в младенчестве без применения антибиотиков, были связаны с повышенным риском развития ожирения в сравнении с детьми, не болевшими инфекциями (ОШ: 1,25; 95% ДИ: 1,20–1,29). Наблюдаемая связь имела доза-зависимый эффект: при увеличении количества эпизодов инфекционных заболеваний увеличивался риск развития ожирения. При этом вероятность развития ожирения у детей, перенёсших инфекции и получавших антибиотики, и детей, перенёсших инфекции, но не получавших антибиотики, не различалась (ОШ: 1,01; 95% ДИ: 0,98–1,04).

DAGs, построенный по заданным в вышеупомянутом примере условиям, представлен на рис. 4А. В верхней

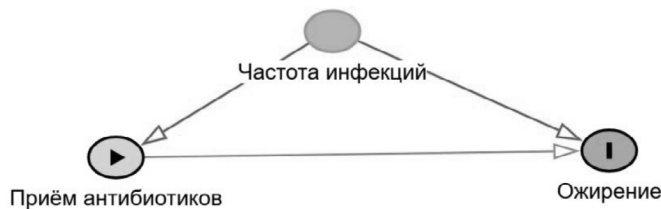
правой части рабочей области интерфейса DAGitty находится вкладка «Causal effect identification» (Оценка причинно-следственной связи), где указывается минимально достаточный набор переменных (Minimal sufficient adjustment sets) для коррекции при оценке общего эффекта (влияния; total effect) воздействия на исход (рис. 4В). В приведённом примере для корректной оценки связи между приёмом антибиотиков и развитием ожирения необходима коррекция на частоту инфекций. Иногда переменную-конфаундер сложно измерить объективно, например, если невозможно оценить количество перенесённых инфекций, коррекция может быть проведена с учётом количества визитов к врачу (прокси-конфаундер), что косвенно может отражать частоту инфекций (рис. 4С). Построив DAGs с учётом прокси-конфаундера, мы видим, что коррекция может быть проведена на конфаундер или прокси-конфаундер (рис. 4D). Коррекция на оба фактора будет избыточной. На рис. 4Е проведена коррекция на прокси-конфаундер. При этом красные стрелки стали чёрными, что значит, что ошибочные (смещённые) пути, проходящие через конфаундер, закрыты. Во вкладке «Causal effect identification» (Оценка причинно-следственной связи) программа сообщает нам, что коррекция проведена верно (рис. 4F).

Медиатор — это переменная, которая является промежуточным звеном в цепочке событий между воздействием и исходом (на пути от воздействия к исходу). Например, при оценке влияния образования (воздействие) на здоровье медиатором может быть уровень дохода, индекс массы тела — при изучении связи между диетой и риском развития сахарного диабета 2-го типа.

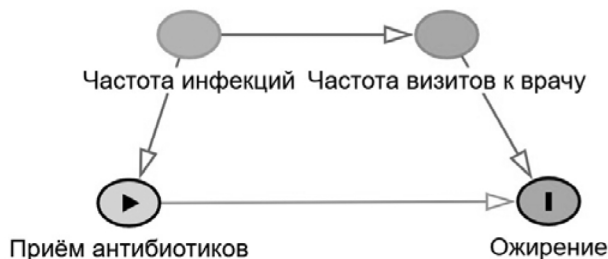
Обычно задачей эпидемиологических исследований является оценка общего эффекта (влияния) воздействия на исход (total effect). Непреднамеренная коррекция на медиатор, например, включение медиатора в регрессионную модель, может привести к занижению общей оценки эффекта воздействия [11, 16]. В то же время задачей исследования может являться оценка непосредственного влияния воздействия на исход, которая определяется понятием «прямой эффект» (direct effect).

Рассмотрим конкретный пример: при оценке связи между приёмом тромболитиков (воздействие) и летальным исходом у пациентов, перенёсших инфаркт миокарда, медиатором могут выступать побочные эффекты лечения. В одном случае это массивное кровотечение, непосредственно приведшее к летальному исходу, в другом — незначительное кровотечение, которое само по себе не стало причиной смерти, но послужило основанием для прекращения терапии. Прекращение приёма препарата, в свою очередь, могло способствовать развитию неблагоприятного исхода. Граф, построенный по заданным условиям, представлен на рис. 5А.

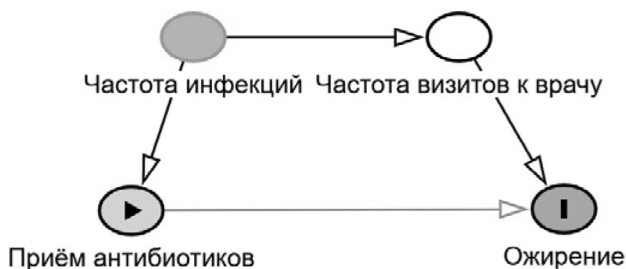
При выборе варианта эффекта — общий эффект (total effect) или прямой эффект (direct effect) — с использованием вкладки «Causal effect identification» (Оценка



A



C



E

☑ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):
Приём антибиотиков
Outcome (Исход): Ожирение
Biasing paths are open
(Ошибочные пути открыты).
Minimal sufficient adjustment sets for estimating the total effect of (минимально достаточный набор переменных для оценки общего эффекта) Приёма антибиотиков on (на) Ожирение:
• Частота инфекций

B

☑ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):
Приём антибиотиков
Outcome (Исход): Ожирение
Biasing paths are open
(Ошибочные пути открыты).
Minimal sufficient adjustment sets for estimating the total effect of (минимально достаточный набор переменных для оценки общего эффекта) Приёма антибиотиков on (на) Ожирение:
• Частота инфекций
• Частота визитов к врачу

D

☑ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):
Приём антибиотиков
Outcome (Исход): Ожирение
Adjusted (Скорректировано):
Частота визитов к врачу
Correctly adjusted
(Скорректировано верно).

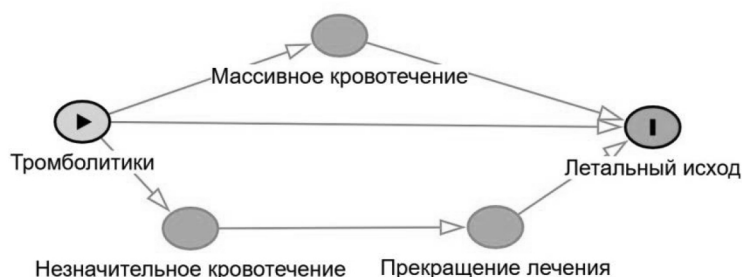
F

Рис. 4. Пример DAGs, отражающего явление конфаундинга.
Fig. 4. Example of a DAGs illustrating the phenomenon of confounding.

причинно-следственной связи) в верхней правой части рабочей области DAGgitty программа сообщает о необходимости коррекции. В нашем примере для оценки общего эффекта воздействия на исход не требуется проведения коррекции на медиаторы (рис. 5B). Однако если для исследователя представляет интерес оценка прямого эффекта препарата, не связанная с возникновением кровотечения, потребуется учёт влияния (коррекция) медиаторов (рис. 5C, 5D). При этом программа предлагает

два варианта коррекции: в первом учитываются переменные «массивное кровотечение» и «незначительное кровотечение», во втором — «массивное кровотечение» и «прекращение лечения». Коррекция с учётом всех трёх переменных будет избыточной, что может неоправданно занижать мощность исследования.

В случае, если интерес представляет оценка именно прямого эффекта, это следует чётко указать в цели исследования, описать возможные медиаторы в методах



A

☒ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):

Тромболитики

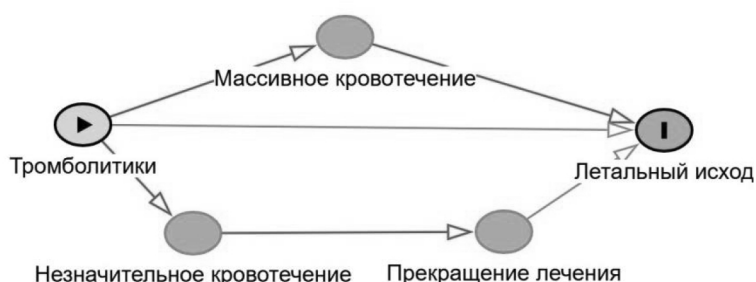
Outcome (Исход): Летальный исход

No open biasing paths
(Нет открытых ошибочных путей).

No adjustment is necessary to estimate the total effect of (не требуется коррекции для оценки эффекта)

Тромболитиков on (на)
Летальный исход

B



C

☒ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (direct effect)
(Коррекция (прямой эффект))

Exposure (Воздействие):

Тромболитики

Outcome (Исход): Летальный исход

Biasing paths are open
(Ошибочные пути открыты).

Minimal sufficient adjustment sets for estimating the direct effect of (минимально достаточный набор переменных для оценки прямого эффекта)

Тромболитиков on (на)
Летальный исход:

- Массивное кровотечение, Незначительное кровотечение
- Массивное кровотечение, Прекращение лечения

D

Рис. 5. Оценка общего и прямого эффекта воздействия на исход.

Fig. 5. Estimating the total and direct effects of exposure on the outcome.

и провести соответствующую коррекцию при выполнении анализа.

Коллайдер (Collider) — это переменная, которая одновременно зависит от двух других переменных, но сама не влияет на эти переменные. В контексте эпидемиологических исследований часто коллайдером является фактор, на который одновременно влияют как изучаемое воздействие, так и изучаемый исход, но при этом сам коллайдер не влияет ни на воздействие, ни на исход.

При отсутствии коррекции на коллайдер путь, проходящий через него, остаётся закрытым. Коррекция на коллайдер, в частности включение коллайдера в регрессионную модель, приводит к открытию данного пути и созданию ложной связи между изучаемыми переменными, которые изначально были независимыми. Такой эффект известен как ошибка коллайдера (collider bias) [2]. Ошибка коллайдера может возникнуть также в результате ошибки отбора участников исследования (selection bias). Типичный пример — ошибка Берксона (Berkson's bias), при которой отбор по переменной-коллайдеру (например, госпитализация) индуцирует искусственную связь между независимыми факторами.

Рассмотрим другой пример. При оценке связи между курением (воздействие) и тяжестью COVID-19 (исход) хронические болезни лёгких являются медиатором, коррекция на который, как мы уже знаем, может привести к недооценке общего эффекта воздействия на исход (рис. 6A). Однако в приведённом примере коррекция на хронические болезни лёгких может также привести к ошибке коллайдера. Помимо курения, на развитие хронических болезней лёгких могут влиять и другие факторы (например, профессиональные вредности, наследственные заболевания, фиброз в результате перенесённой ранее инфекции и т.д.), которые также могут повышать риск тяжёлого течения COVID-19. В такой ситуации коррекция на хронические болезни лёгких приводит к ошибке коллайдера, создавая ложную связь между курением и другими причинами хронических заболеваний лёгких. Это искажает истинную ассоциацию между курением и тяжестью COVID-19 (рис. 6C). В исследовании, изучавшем факторы риска тяжёлого течения коронавирусной инфекции в одномерном анализе без учёта других факторов, курение связано с повышенным риском тяжёлого течения COVID-19 (ОШ: 1,14; 95% ДИ: 1,05–1,23) [17]. Однако после включения хронических болезней лёгких в регрессионную модель (после коррекции) направление связи изменилось (ОШ: 0,89; 95% ДИ: 0,82–0,97), что может быть результатом ошибки коллайдера. Такой результат следует интерпретировать не как независимый эффект курения на тяжесть инфекции, а как совместный эффект (joint effect): курение само по себе может повышать риск тяжёлого COVID-19, но не в такой степени, как другие причины хронических болезней лёгких. Визуализация DAGs помогает избежать такой ошибки: программа не предлагает корректировать

модель на коллайдер (рис. 6B), а при попытке включения коллайдера в модель сообщает об ошибке (рис. 6D).

Рассмотрим более сложный пример построения DAGs (рис. 7A). Представленный граф иллюстрирует структуру связей между изучаемыми факторами при оценке влияния ожирения (воздействие) на тяжесть COVID-19 (исход). DAGs построен на основании совокупности эпидемиологических и клинических исследований, посвящённых факторам риска тяжёлого течения COVID-19 [17–19].

В структуру графа включены все факторы, которые, согласно публикациям и представлениям исследователя, связаны с воздействием и/или исходом. Факторы (переменные), сведения о которых отсутствуют, окрашены в серый цвет.

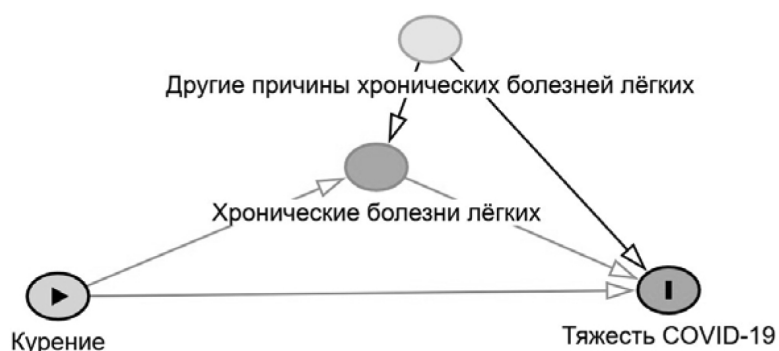
Пол, возраст, уровень образования, злоупотребление алкоголем, курение могут быть связаны как с ожирением, так и с тяжестью COVID-19, то есть являются конфаундерами [19].

Ожирение ассоциировано с хроническим низкоуровневым системным воспалением, повышенным риском цитокинового шторма и гиперкоагуляцией, что способствует более тяжёлому и осложнённому течению COVID-19 [18]. Также наличие ожирения связано с атерогенной дислипидемией, гипертонией и сахарным диабетом 2-го типа — состояниями, которые увеличивают риск тяжёлого течения заболевания [20]. Гипертония и диабет рассматриваются как медиаторы, то есть переменные, лежащие на причинном пути от воздействия (ожирения) к исходу (тяжесть COVID-19). Поскольку целью исследования является оценка общего эффекта ожирения (total effect), включение этих медиаторов в модель в качестве ковариат может привести к смещению из-за избыточной коррекции (overadjustment) и, соответственно, к недооценке общего эффекта.

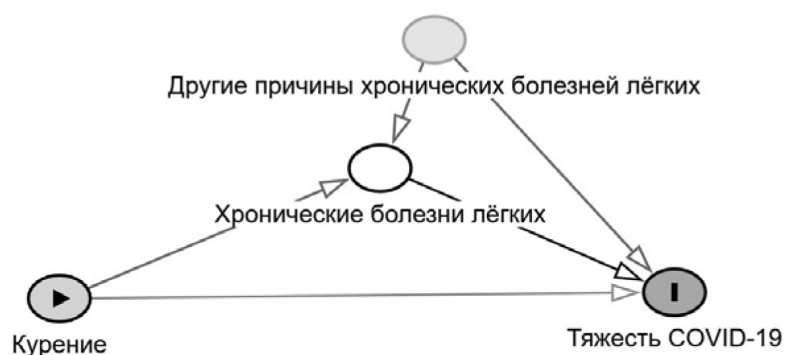
Обращение за медицинской помощью может зависеть как от наличия у пациента определённых состояний или заболеваний (в том числе ожирения), так и от исхода (тяжести заболевания), следовательно, это коллайдер. Включение данной переменной в регрессионную модель может привести к ошибке коллайдера (collider bias) [21].

Для оценки общего эффекта программа предлагает коррекцию на три из пяти выявленных конфаундеров: возраст, злоупотребление алкоголем, курение (рис. 7B). Поскольку пол и уровень образования являются прокси-конфаундерами по отношению к курению и употреблению алкоголя, коррекция на них является избыточной. Рис. 7C отражает граф после коррекции на возраст, злоупотребление алкоголем, курение. Мы видим, что все ошибочные (смещённые) пути, обозначенные красными стрелками на рис. 7A, закрыты (рис. 7C). Следовательно, коррекция проведена верно (рис. 7D).

Таким образом, при проведении эпидемиологических исследований коррекцию на этапе статистической обработки, например, при регрессионном анализе, необходимо выполнять с учётом роли переменных (ковариат)



A



C

☑ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):

Курение

Outcome (Исход): Тяжесть COVID-19

No open biasing paths

(Нет открытых ошибочных путей).

No adjustment is necessary to estimate the total effect of (не требуется коррекции для оценки эффекта) Курения on (на) Тяжесть COVID-19

B

☑ Causal effect identification
(Оценка причинно-следственной связи)

Adjustment (total effect)
(Коррекция (общий эффект))

Exposure (Воздействие):

Курение

Outcome (Исход): Тяжесть COVID-19

Adjusted (Скорректировано):

Хронические болезни лёгких

Incorrectly adjusted (Неверно скорректировано).

No adjustment sets found.

(Переменных для коррекции не найдено / коррекции не требуется).

D

Рис. 6. Пример DAGs, отражающего ошибку коллайдера.

Fig. 6. Example of a DAG illustrating collider bias.

по отношению к воздействию и исходу. В зависимости от типа переменной коррекция может как устранить возможное искажение (конфаундер), так и привести к недооценке (медиатор) или искажению изучаемой связи (коллайдер). При построении DAG в граф следует включить все переменные, которые по данным литературы или исходя из гипотезы и допущений исследователя связаны с изучаемыми воздействием и/или исходом и важны для оценки связи, включая те, для которых отсутствуют прямые измерения. Точность определения набора переменных, необходимых для коррекции, и точность оценки искомой связи будет зависеть от того, насколько точно DAG отражает реально существующие связи между переменными. Следует также учитывать ограничения дизайна исследования,

например, в поперечных исследованиях связей/ассоциаций не всегда ясны причинно-следственные отношения между анализируемыми факторами (сложно однозначно определить направление стрелок) в связи с отсутствием данных о последовательности изучаемых событий (факт воздействия и наступления исхода) во времени.

Минимальный и достаточный набор переменных для коррекции определяется как набор переменных, который устраняет искажения в оценке связи, закрывая все ложные пути и оставив все причинно-следственные пути открытыми [7]. Если одна или несколько переменных недоступны, следует учесть максимально возможное число переменных, необходимых для коррекции по результатам построения DAGs [22]. Отсутствие данных о переменных, на которые требуется коррекция, следует указать

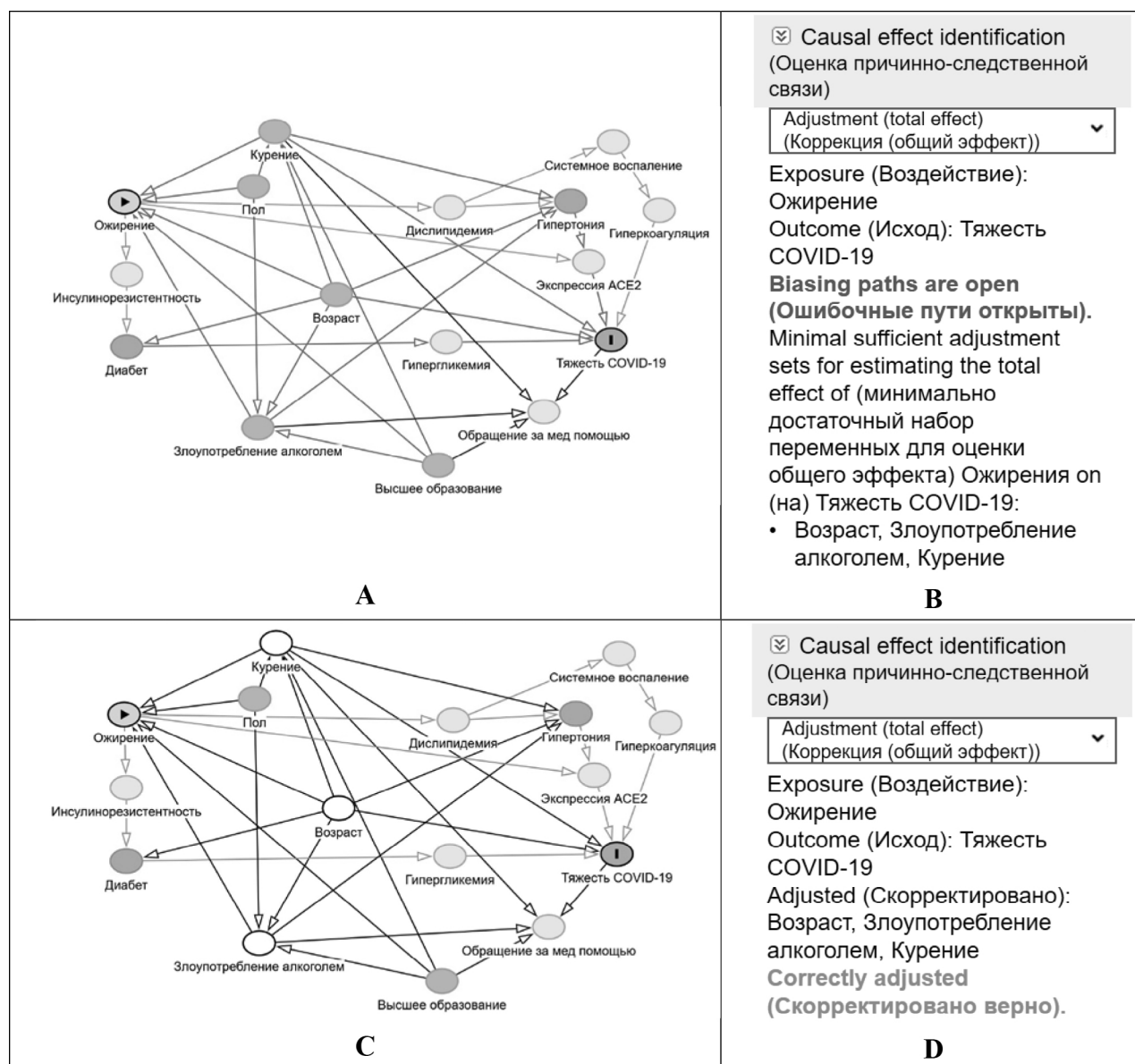


Рис. 7. Пример построения DAGs для определения переменных, которые необходимо учитывать при оценке влияния ожирения на тяжесть COVID-19.

Fig. 7. Example of DAG construction to identify variables that must be considered when assessing the impact of obesity on COVID-19 severity.

как ограничение исследования, так как это может, например, свидетельствовать о потенциальном резидуальном (остаточном) конфаундинге.

Для обеспечения прозрачности процесса отбора переменных для анализа, сравнимости и интерпретируемости результатов DAGs, использованный для отбора переменных при построении регрессионной модели для оценки влияния изучаемого воздействия (воздействий) на исход, рекомендуется размещать в основном тексте статьи при описании методологии исследования в разделе «Методы исследования» или в дополнительных материалах. В описание графа, особенно для сложных DAGs, следует добавлять текст с объяснением предположений и допущений исследователя, литературными ссылками, объясняющими наличие связи между переменными.

ЗАКЛЮЧЕНИЕ

Таким образом, можно выделить основные рекомендации по составлению, интерпретации и использованию DAGs в эпидемиологических исследованиях.

1. Построение DAGs целесообразно в исследованиях, которые в качестве гипотезы используют предположение о наличии причинно-следственной связи между воздействием и исходом.

2. Построение DAGs проводится до начала сбора данных или перед проведением статистического анализа (при вторичном анализе ранее собранных данных).

3. Воздействие, исход, изучаемая связь должны быть чётко определены и отражены в цели исследования.

4. В DAGs должны быть представлены все основные переменные, потенциально связанные с изучаемыми

воздействием и/или исходом, включая те, для которых отсутствуют прямые измерения.

5. Выбор переменных, включённых в DAGs, и связей между ними базируется на результатах анализа результатов ранее проведённых исследований.

6. DAGs должны быть визуально ясны и читаемы.

7. DAGs, используемые при отборе переменных для анализа, могут быть размещены в основном тексте статьи в разделе «Методы исследования» или в дополнительных материалах; в описании графа должны присутствовать объяснения наличия связи между переменными со ссылками на ранее проведённые исследования, использованные при подготовке DAGs.

8. Если одна или несколько переменных, необходимых для корректной оценки связи, недоступны, при проведении анализа следует учесть максимально возможное число переменных, необходимых для коррекции по данным построенных DAGs.

9. Отсутствие в исследовании данных по переменным, учёт которых при анализе требуется на основании DAGs, следует указать как ограничение исследования.

10. Построение DAGs возможно с использованием бесплатно распространяемого программного обеспечения.

Авторы выражают надежду, что представленная информация окажется полезной для исследователей и станет стимулом к более широкому применению современных подходов к анализу (причинно-следственных) связей. Использование DAGs для обоснованного выбора переменных для включения в многомерные математические модели будет способствовать повышению точности, надёжности и воспроизводимости результатов в биомедицинской науке.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Вклад авторов. Е.А. Кригер — концепция и структура обзора, поиск и анализ литературных источников, написание основного текста, подготовка иллюстраций, редактирование рукописи; В.А. Постоев — критический анализ материала, участие в написании текста, редактирование и научное руководство; А.В. Кудрявцев — экспертная оценка содержания, редактирование рукописи; Т.Н. Унгуриану — оптимизация способов представления информации, редактирование рукописи; А.М. Гржибовский — обобщение методологических подходов, редактирование рукописи. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы

внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Этическая экспертиза. Статья носит обзорный методологический характер. В связи с этим экспертиза комитетом по этике не проводилась.

Источники финансирования. Отсутствуют.

Раскрытие интересов. Авторы заявляют об отсутствии отношений, деятельности и интересов за последние три года, связанных с третьими лицами (коммерческими и некоммерческими), интересы которых могут быть затронуты содержанием статьи.

Оригинальность. При создании настоящей работы авторы не использовали ранее опубликованные сведения (текст, иллюстрации, данные).

Доступ к данным. Редакционная политика в отношении совместного использования данных к настоящей работе не применима, новые данные не собирали и не создавали.

Генеративный искусственный интеллект. При создании настоящей статьи технологии генеративного искусственного интеллекта не использовали.

Рассмотрение и рецензирование. Настоящая работа подана в журнал в инициативном порядке и рассмотрена по обычной процедуре. В рецензировании участвовали два внешних рецензента, член редакционной коллегии и научный редактор издания.

ADDITIONAL INFORMATION

Author contributions: E.A. Krieger: conceptualization, investigation, writing—original draft, visualization, writing—review & editing; V.A. Postoev: formal analysis, writing—original draft, writing—review & editing, supervision; A.V. Kudryavtsev: validation, writing—review & editing; T.N. Unguryanu: visualization, writing—review & editing; A.M. Grijbovski: methodology, writing—review & editing. All the authors confirm that their authorship meets the ICMJE criteria (all authors made substantial contributions to the conceptualization, investigation, and manuscript preparation, and reviewed and approved the final version prior to publication).

Ethics approval: This article is a methodological review. Therefore, ethical approval by an ethics committee was not required.

Funding sources: No funding.

Disclosure of interests: The authors have no relationships, activities, or interests for the last three years related to for-profit or not-for-profit third parties whose interests may be affected by the content of the article.

Statement of originality: No previously published material (text, images, or data) was used in this work.

Data availability statement: The editorial policy regarding data sharing does not apply to this work, as no new data was collected or created.

Generative AI: No generative artificial intelligence technologies were used to prepare this article.

Provenance and peer-review: This paper was submitted unsolicited and reviewed following the standard procedure. The peer review process involved two external reviewers, a member of the editorial board, and the in-house scientific editor.

СПИСОК ЛИТЕРАТУРЫ | REFERENCES

1. Mitkin NA, Drachev SN, Krieger EA, et al. Sample size calculation for cross-sectional studies. *Ekologiya cheloveka (Human Ecology)*. 2023;30(7):509–522. doi: 10.17816/humeco569406 EDN: LOEJVM
2. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37–48.
3. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:70. doi: 10.1186/1471-2288-8-70
4. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34(3):211–219. doi: 10.1007/s10654-019-00494-6
5. Foraita R, Spallek J, Zeeb H. Directed acyclic graphs. In: Ahrens W, Pigeot I, editors. *Handbook of Epidemiology*. New York: Springer; 2014. P. 1481–1517. doi: 10.1007/978-0-387-09834-0_65
6. Kornaropoulos EM, Tollis IG. DAG View: an approach for visualizing large graphs. In: Didimo W, Patrignani M, editors. *Graph Drawing. GD 2012. Lecture Notes in Computer Science*. Vol. 7704. Berlin, Germany: Springer; 2013. P. 499–510. doi: 10.1007/978-3-642-36763-2_44
7. Textor J, Zander BVD, Gilthorpe MS, et al. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol*. 2016;45(6):1887–1894. doi: 10.1093/ije/dyw341
8. Porta MS, Greenland S, Hernanet M, et al. *A dictionary of epidemiology*. 6th ed. New York: Oxford University Press; 2014. 343 p.
9. Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. 4th ed. Sudbury: Jones & Bartlett Learning, LLC; 2018. 489 p.

10. Bursac Z, Gauss CH, Williams DK, et al. Purposeful selection of variables in logistic regression. *Source Code Biol Med.* 2008;3:17. doi: 10.1186/1751-0473-3-17
11. Textor J, Liśkiewicz M. Adjustment criteria in causal diagrams: an algorithmic perspective. In: *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*. Barcelona, Spain: AUAI Press; 2011. doi: 10.48550/arXiv.1202.3764
12. Lee PH. Is a cutoff of 10% appropriate for the change-in-estimate criterion of confounder identification? *J Epidemiol.* 2014;24(2):161–167. doi: 10.2188/jea.je20130062
13. Heinze G, Wallisch C, Dunkler D. Variable selection — A review and recommendations for the practicing statistician. *Biom J.* 2018;60(3):431–449. doi: 10.1002/bimj.201700067
14. Bailey LC, Forrest CB, Zhang P, et al. Association of antibiotics in infancy with early childhood obesity. *JAMA Pediatr.* 2014;168(11):1063–1069. doi: 10.1001/jamapediatrics.2014.1539
15. Li DK, Chen H, Ferber J, et al. Infection and antibiotic use in infancy and risk of childhood obesity: a longitudinal birth cohort study. *Lancet Diabetes Endocrinol.* 2017;5(1):18–25. doi: 10.1016/S2213-8587(16)30281-9
16. Dekkers OM, Laugesen K, Groenwold RHH. Directed acyclic graphs in clinical research. *Eur J Endocrinol.* 2024;190(4):E5–E7.
17. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature.* 2020;584(7821):430–436. doi: 10.1038/s41586-020-2521-4
18. Sattar N, McInnes IB, McMurray JJV. Obesity is a risk factor for severe COVID-19 infection: multiple potential mechanisms. *Circulation.* 2020;142(1):4–6. doi: 10.1161/CIRCULATIONAHA.120.047659
19. Hamer M, Kivimäki M, Gale CR, Batty GD. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: A community-based cohort study of 387,109 adults in UK. *Brain Behav Immun.* 2020;87:184–187. doi: 10.1016/j.bbi.2020.05.059
20. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med.* 2020;8(4):e21. doi: 10.1016/S2213-2600(20)30116-8
21. Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun.* 2020;11(1):5749. doi: 10.1038/s41467-020-19478-2
22. Lash MP, Fink AK. *Applying quantitative bias analysis to epidemiologic data*. New York: Springer; 2011. 192 p.

ОБ АВТОРАХ

***Кригер Екатерина Анатольевна**, канд. мед. наук, PhD, доцент;
адрес: Россия, 163069, Архангельск, пр-т Троицкий, д. 51;
ORCID: 0000-0001-5179-5737;
eLibrary SPIN: 2686-7226;
e-mail: kate-krieger@mail.ru

Постоев Виталий Александрович, канд. мед. наук, PhD, доцент;
ORCID: 0000-0003-4982-4169;
eLibrary SPIN: 6070-2486;
e-mail: v.postoev@nsmu.ru

Кудрявцев Александр Валерьевич, PhD;
ORCID: 0000-0001-8902-8947;
eLibrary SPIN: 9296-2930;
e-mail: alex.v.kudryavtsev@yandex.ru

Унгуряну Татьяна Николаевна, д-р мед. наук, PhD, доцент;
ORCID: 0000-0001-8936-7324;
eLibrary SPIN: 7358-1674;
e-mail: unguryanu_tn@mail.ru

Гржибовский Андрей Мечиславович, доктор медицины, PhD;
ORCID: 0000-0002-5464-0498;
eLibrary SPIN: 5118-0081;
e-mail: a.grjibovski@yandex.ru

AUTHORS' INFO

***Ekaterina A. Krieger**, MD, Cand. Sci. (Medicine), PhD, Associate Professor;
address: 51 Troitskiy ave, Arkhangelsk, Russia, 163069;
ORCID: 0000-0001-5179-5737;
eLibrary SPIN: 2686-7226;
e-mail: kate-krieger@mail.ru

Vitaly A. Postoev, MD, Cand. Sci. (Medicine), PhD, Associate Professor;
ORCID: 0000-0003-4982-4169;
eLibrary SPIN: 6070-2486;
e-mail: v.postoev@nsmu.ru

Alexander V. Kudryavtsev, PhD;
ORCID: 0000-0001-8902-8947;
eLibrary SPIN: 9296-2930;
e-mail: alex.v.kudryavtsev@yandex.ru

Tatiana N. Unguryanu MD, Dr. Sci. (Medicine), PhD, Associate Professor;
ORCID: 0000-0001-8936-7324;
eLibrary SPIN: 7358-1674;
e-mail: unguryanu_tn@mail.ru

Andrej M. Grjibovski, MD, PhD;
ORCID: 0000-0002-5464-0498;
eLibrary SPIN: 5118-0081;
e-mail: a.grjibovski@yandex.ru

* Автор, ответственный за переписку / Corresponding author