

МЕТОДОЛОГИЯ
НАУЧНОГО ПОИСКА

УДК 167:61:573.01:616.036.22

КРИТЕРИИ ПРИЧИННОСТИ В МЕДИКО-БИОЛОГИЧЕСКИХ
ДИСЦИПЛИНАХ: ИСТОРИЯ, СУЩНОСТЬ И РАДИАЦИОННЫЙ АСПЕКТ.
СООБЩЕНИЕ 4, ЧАСТЬ 2: ИЕРАРХИЯ КРИТЕРИЕВ, ИХ КРИТИКА
И ИНЫЕ МЕТОДЫ УСТАНОВЛЕНИЯ ПРИЧИННОСТИ[#]

© 2022 г. А. Н. Котеров^{1,*}, Л. Н. Ушенкова¹

¹ ГНЦ РФ – Федеральный медицинский биофизический центр им. А.И. Бурназяна ФМБА России, Москва, Россия

*E-mail: govorilga@inbox.ru

Поступила в редакцию 12.05.2021 г.

После доработки 10.09.2021 г.

Принята к публикации 22.12.2021 г.

В части 2 сообщения 4 рассмотрены три темы. Во-первых – попытки систематизации, взвешивания и построения иерархии (“рангов”) для критериев причинности. Использовались две методологии: обращение к конкретным разработкам различных исследователей и оценка путем определения частоты встречаемости того или иного критерия в тех работах, в которых они использовались как способ установления каузальности эффектов. Если не считать априорно лидирующего критерия “Временная зависимость”, на первые два места следует поставить “Силу связи” и “Постоянство ассоциации” (последовательность их может быть и обратной), т.е. индуктивные пункты. Третье место занимает “Эксперимент контрафактический” (“природный эксперимент”), который, однако, доступен в эпидемиологии и в экологии достаточно редко. При его отсутствии на третье место выходит вновь индуктивный критерий “Биологический градиент” (зависимость “доза–эффект”). В областях экологии и экотоксикологии, где редко доступны экспериментальные или наблюдательные исследования на людях, ранги критериев специфичны и на первом месте находится “Биологическое правдоподобие”.

Во-вторых, представлена критика каузального подхода в наблюдательных дисциплинах, основанного на критериях причинности, начиная от отсутствия доказательности самого индуктивного метода. Указывается на “мозаичность”, мультидисциплинарность и субъективизм критериев, равно как и их неабсолютность (“обременены оговорками и исключениями”). Отмечается, однако, что подобные вопросы можно поставить почти перед любыми подходами в области естественнонаучных дисциплин. Существует мнение, что критерии причинности, как бы “подтверждающие эффекты”, надо оставить только для принятия решений в области здравоохранения, а не для строгой доказательности в области “чистой науки” с ее фальсификацией гипотез. Вряд ли этот подход может быть принят для эпидемиологии.

В-третьих, рассмотрены модели оценки причинности эффектов в эпидемиологии вне каузальных критериев. Самой известной является модель достаточной компонентной причины K.J. Rothman (“пирог причинности”), подразумевающая детерминизм эпидемиологических эффектов, в отличие от позиции практической эпидемиологии и здравоохранения, в которой причинность определяется как вероятностная. Второй моделью называется контрафактическая или “модель потенциальных исходов”, третьей – направленные ациклические графы или причинные диаграммы (DAG), а четвертой – модель структурных уравнений. Еще обнаружен подход C.V. Phillips, заключающийся в количественной оценке ошибок (неопределенностей) на всех этапах эпидемиологического исследования с моделированием методом Монте-Карло. Сделан вывод, что названные модели оказываются или качественными и иллюстративными, или представляют из себя лишь графический либо математический аппарат, служащий все тому же исходному подходу, основанному на причинных критериях.

Ключевые слова: критерии причинности, критерии A.V. Hill, модель достаточной компонентной причины, контрафактическая модель, модель направленных ациклических графов, модель структурных уравнений, эпидемиология

DOI: 10.31857/S0869803122040051

Цикл наших сообщений (всего четыре; два в нескольких частях), вместе с сопутствующими

статьями-преамбулами (всего пять), посвящен истории появления, разработке, использованию, актуальности и ограничениям критериев причинности в наблюдательных дисциплинах [1–10].

[#] Публикуется в авторской редакции.

Наиболее известны девять руководящих принципов (“viewpoints”) Хилла (A.V. Hill) [11], восемь из которых этот авторитетный английский статистик в области медицины только собрал воедино, взяв у других авторов [2]. Тем не менее теперь критерии причинности в эпидемиологии разных направлений называются почти всегда “критериями Хилла” или “руководящими принципами Хилла” [2, 3, 5–8, 10].

Рассмотрев общие модели и определения причинности в философии, медицине и эпидемиологии [1], равно как и все причинные критерии Хилла по отдельности [3–9], ранее в историческом обзоре [2] мы изложили истоки их появления в эпидемиологии, перечислив истинных пионеров – авторов “до Хилла”. Но и “после Хилла” рядом авторов и организаций продолжались попытки совершенствования или модификации методов оценки причинности ассоциаций и эффектов в медико-биологических дисциплинах. Этой теме была посвящена часть 1 настоящего Сообщения 4, самым важным в которой являются принципы оценки индивидуальной причинности в профессиональной медицине и оценка каузальных эффектов в экологических дисциплинах. Все из них имеют отношение к критериям Хилла [10].

В части 2 Сообщения 4 рассмотрены оставшиеся пока без подробного освещения следующие проблемы:

1. Попытки выработать градации значимости тех или иных критериев, определить их весовую, равно как выполнить их ранговую рубрификацию с позиции научной философии (концептуально) или в плане специальных отражений доказательности.

2. Критика методов оценки причинности, основанных на критериях и руководящих принципах, равно как вопрос о том, насколько она обоснована и, главное, конструктивна в практическом плане.

3. Другие методологии оценки причинности, основанные не на критериях (“руководящих принципах”, “пунктах” и т.д.), а на иных моделях (“causal graph” и пр.) [12–27]. В нашу задачу не входит особенно подробное рассмотрение этих, как правило, статистических, математических, вероятностных и графических подходов, которые, судя по всему, применяются нечасто. Но ознакомиться с ними – целесообразно.

Частью 3 настоящего Сообщения 4 запланировано закончить наш цикл, посвятив его (часть) широте использования критериев Хилла в той или иной форме в разных дисциплинах и разными международными и имеющими международный авторитет организациями, разрабатывающими рекомендации по “Весу свидетельств” – Weight of Evidence (WoE) [28]. Организации, оценивающими риски и принимающими реше-

ния в области здравоохранения и безопасности населения, включая радиационные сферы. Этот момент несколько затрагивался нами ранее [1, 2, 6, 9, 10], но особенно – в части 1 Сообщения 4, где были приведены данные по использованию критериев и постулатов причинности в экологии биоты, экоэпидемиологии и экотоксикологии [10].

ГРАДАЦИИ, ИЕРАРХИИ И ВЗВЕШИВАНИЕ КРИТЕРИЕВ ПРИЧИННОСТИ

Классический список критериев Хилла из его основополагающей работы 1965 г. [11], который неоднократно приводился нами ранее [2, 8–10], следующий (понадобится далее):

- 1) сила связи (Strength of the association);
- 2) постоянство ассоциации (Consistency of the association);
- 3) специфичность ассоциации (Specificity of the association);
- 4) временная зависимость (Temporality);
- 5) биологический градиент (Biological gradient) или зависимость “доза – эффект”;
- 6) биологическое правдоподобие (Biological plausibility);
- 7) согласованность с текущими фактами и теоретическими знаниями (Coherence);
- 8) эксперимент (Experiment);
- 9) аналогия (Analogy).

Восемь из девяти “критериев” – не правила и не алгоритмы установления причинности

Для термина “критерий” на русском и английском языках находится много определений, среди которых, помимо “мерила оценки, суждения” (словарь С.И. Ожегова и Н.Ю. Шведовой от 1992 г.), есть и “правило принятия решения по оценке чего-либо” (decision rule) [29]. То есть подразумевается жесткий принцип, основа для суждения, которое делает последнее как бы стопроцентным. Но почти все критерии Хилла таковыми не являются, о чем нами не раз упоминалось ранее [3, 5, 6, 8]. Заметим, что большинство предшественников A.V. Hill, за исключением Lilienfeld A.M., 1959 [30], не использовали термин “критерии” вплоть до 1963–1964 гг. (Сообщение Главного врача (Surgeon General) США о последствиях курения от 1964 г. [31] и соответствующий его драфт о каузации Stallones R.A., 1963 [32]). Список терминов для причинных подходов-пунктов “до Хилла” приводился нами ранее [2]: это и “меры предосторожности” (precautions), и “характеристики”, и “факторы”, и “указатели” (guideposts), и “руководящие принципы” (guidelines), и “пункты” (points). В одном случае, правда, имелось на-

именование “постулаты” (Wynder E.L., 1956; Wynder E.L., Day E., 1961; см. в [2]).

Сам А.В. Hill использовал термин “точки зрения” или “положения” (“viewpoints”), приведя рассуждения [11], которые затем цитировались аутентично много раз [13, 15, 33–40], не говоря уже о просто упоминаниях:

“Есть девять разных положений (viewpoints), все из которых следует изучить применительно к ассоциации, прежде чем вводить причинность. То, во что я не верю, – и это установлено, – так это то, что мы можем с пользой применить некоторые жесткие и оперативные (hard-and-fast) правила доказательности, обязанные соблюдаться, если мы определяем причину эффекта. Ни одна из моих девяти точек зрения не может дать неоспоримые доказательства “за” или “против” гипотезы причинно-следственной связи, и ни одна из них не может требоваться в качестве *sine qua non* [обязательного условия]. То, что они могут в большей или меньшей степени делать, так это помочь нам решить основной вопрос – есть ли другой способ объяснить множество фактов, находящихся в нашем распоряжении, имеется ли какой-либо другой равнозначный или более вероятный ответ, чем причина и следствие?”¹ (список примечаний идет после основного текста).

Как сам А.В. Hill (нередко “вольный художник” [2]²), так и его цитирующие, не заметили за лет 50 ошибки, ибо один пункт – “Временная зависимость”, понятно, является истинным критерием (это отмечалось еще Д. Юмом и Дж. Миллем [5]). Его нарушение (фальсификация) сразу отменяет причинность, и никакие иные руководящие принципы уже не нужны [5, 40–44]³. К сожалению, что рассматривалось нами ранее [5], не все столь просто: в эпидемиологии и экологии, в реконструктивных дисциплинах (история, археология палеонтология и пр.) и даже в обыденной жизни порой затруднительно установить временной порядок.

Рассуждения о том, что руководящие принципы Хилла не могут использоваться в качестве строго количественного контрольного списка (checklist), а скорее применяются для определения WoE (веса свидетельств) или силы доказательств применительно к заключению о причинно-следственной связи [45], встречаются практически в каждой публикации на тему (например, [13, 24, 33, 34, 44, 46–60]). Более того, если бы критерии были действительно *контрольным списком* “критериев”, то дело в плане доказательности эффектов обстояло бы плохо. В Weed D.L., 2004 [50] сказано, что если для отклонения причинной гипотезы требуется несоблюдение какого-либо единственного критерия, любого из списка, то “аксиоматично”, что чем больше критериев, тем менее вероятно, что вывод можно будет сделать.

Значимость методологии, основанной на причинных критериях

Если критерии Хилла – это не “критерии”, то как можно судить о том, насколько они близки к последним? Можно ли заявлять при их применении о достаточном уровне дизайна исследования? Ответить на данный вопрос, понятно, затруднительно, но уместно привести цитаты из весомых источников, на которые можно опереться при необходимости:

- “Эти критерии предлагают полезный способ оценки доступной информации” (U.S. Environmental Protection Agency – USEPA; 1998) [61]⁴.

- “Твердый консенсус среди эпидемиологов заключается в том, что выводы о причинно-следственной связи не следует делать... до тех пор, пока не будет рассмотрен ряд критериев” (в судах США; 2004) [62]⁵.

- “Критерии Хилла регулярно цитируются как авторитетные положения (authoritative statements) надлежащего метода оценки совокупности этиологических доказательств” (исходно (2004) [63], затем (2009) [64])⁶.

- “Критерии Бредфорда Хилла являются наилучшими из доступных критериев для причинного вывода”. “...критерии Бредфорда Хилла... по-прежнему наиболее значимые критерии, которые следует использовать при установлении причинно-следственной связи”. “...критерии Бредфорда Хилла все еще остаются ключевыми компонентами причинно-следственной связи” (2009) [55]⁷.

- “Мы считаем, что рекомендации Бредфорда Хилла представляют собой полезный инструмент в их нынешнем виде” (2009) [37]⁸.

- “Предложены различные модификации этого списка, и многие из элементов остаются краеугольными камнями суждения о том, действительно ли воздействие вызывает заболевание, или является ли вмешательство эффективным для предотвращения или терапии патологии” (2011) [56]⁹.

- “Еще одна выгода использования критериев Хилла заключалась в том, что они обеспечивали основу, в рамках которой можно было бы предусмотреть возможность получения неполных наборов данных и отсутствующей информации” (2011) [38]¹⁰.

- “Руководящие принципы, предложенные Бредфордом Хиллом, часто используются для оценки уверенности (confidence) в том, что фактор риска (например, радиационное воздействие) может привести к увеличению числа случаев заболеваний среди населения (например, раком)” (НКДАР-2012) [65]¹¹.

- “Критерии Бредфорда Хилла остаются одной из наиболее цитируемых концепций при ис-

следованиях в области здравоохранения и до сих пор служат надежными инструментами для содействия установлению причинности” (2015) [66]¹².

- “...они выдержали испытание временем как ценный способ оценки силы причинного следствия, когда наличие ассоциации продемонстрировано” (2019) [60]¹³.

- “Несмотря на возобновившийся в последние несколько лет интерес к другим моделям причинности..., руководящие принципы Хилла остаются краеугольным камнем причинно-следственной связи для практического эпидемиолога и эксперта по политике здравоохранения” (2019) [27]¹⁴.

- “Несмотря на быстрое развитие теоретической эпидемиологии, они [критерии Хилла] остаются, по крайней мере, ориентиром для причинно-следственного мышления в наблюдательных комитетах (review committees) и для лиц, принимающих решения” (2019) [67]¹⁵.

Многие авторы не считают возможным вводить градации, иерархии, вес и ранги для отдельных критериев причинности

Известный историк эпидемиологии и доказательной медицины (ЕВМ), в частности причинных критериев, А. Morabia в главе пособия по эпидемиологии от 2014 г. приписал этот подход самому А.В. Hill: “Хилл также подчеркнул, что бессмысленно объединять критерии по баллам (scores), поскольку ассоциация не будет с большей вероятностью причинной, если она соответствует большему числу критериев” [68]¹⁶ (возможно, спутано с вышеприведенными высказываниями из Weed D.L., 2004 [50]). Мы не обнаружили ничего подобного ни в каких публикациях А.В. Hill, хотя его выдержавшие 12 изданий “Principles of Medical Statistics” (1937–1991) нам доступны только в 6 издании (1955), в переводе 1958 г. [69], когда еще никаких “критериев Хилла” не имелось [2]. Однако К.В. Goodman и С.В. Phillips (постоянные авторы статей о каузальности в разных энциклопедиях [18, 55]) отмечали [18], что А.В. Hill не предложил никакой иерархии значимости критериев и ничего не менял в своих “viewpoints” в четырех изданиях “Principles of Medical Statistics” с 1971 по 1991 г.¹⁷ (А.В. Hill умер в 1991 г. [2]).

Правильнее, судя по источникам, второе утверждение — А.В. Hill был, как всегда [11, 70], осторожен.

Согласно Susser M., 1986a [41], который допускает некоторую иерархию, тем не менее критерии суждения почти всегда должны использоваться совместно, как взаимодополняющие. Сходные рассуждения, что значимость одних критериев не умаляет значимости других, и что необходим не-

кий их баланс, имеется и у иных авторов [24, 57, 71–75].

Градации критериев Хилла

Есть ряд работ, в которых предложена градация или рубрификация критериев причинности по какому-либо принципу или признаку, но все они разные. И все остались как бы втуне — не встретилось ни одного обзора, монографии или пособия, где бы это обсуждалось в целом (а не как конкретная схема автора). Систематизация данного материала нелегко поддается логике, поэтому последовательность изложения далее — по нашим соображениям.

Индуктивные и дедуктивные критерии

Такое подразделение есть в Weed D.L., 1988 [76], причем поддержка дедуктивной гипотезы тем или иным критерием зависит от того, какой тип причинности рассматривается, моно, компонентная (подробнее см. в [1]) или мультистадийная.

Непосредственно (direct) дедуктивные критерии: — “Постоянство ассоциации” и “Временная зависимость” (при моно причине), “Временная зависимость” (при компонентной причине), “Сила ассоциации”, “Постоянство ассоциации”, “Временная зависимость”, “Эксперимент контрафактический” (“Preventability” [76]) и “Биологический градиент” (при мультистадийной причинности).

Недедуктивные, пока не изменена гипотеза (not deducible unless hypothesis is altered): “Эксперимент контрафактический”, “Сила ассоциации”, “Специфичность” и “Биологический градиент” (при моно причине), “Постоянство ассоциации”, “Эксперимент контрафактический”, “Сила ассоциации”, “Специфичность” и “Биологический градиент” (при компонентной причине).

Недедуктивные в принципе (not deducible): “Правдоподобие” [биологическое], “Согласованность” с текущими фактами и теоретическими знаниями и “Аналогия” (при моно и компонентной причинах), “Специфичность”, “Правдоподобие”, “Согласованность” с текущими фактами и теоретическими знаниями и “Аналогия” (при мультистадийной причинности).

Логика отнесения того или иного критерия к дедуктивным подразделам в [76], как видим, зависит от типа причинности.

В публикации Vineis P., 1991 [71] указывается, что критерии по списку Хилла № 1–3, 7–9 попадают в категории исчисляющей (enumerative) индукции, элиминативной (eliminative) индукции, дедукции и аналогии, причем все эти категории необходимы для оценки причинно-следственной связи. Пункты же № 4–6 относятся к механисти-

ческой теории. Согласно [71], среди критериев не существует иерархии, хотя дедуктивный тип дизайна исследования является предварительным для любой оценки.

Исчисляющая индукция – это верификационная часть эпидемиологических рассуждений, которая заключается в сборе и подсчете положительных примеров связи между воздействием и эффектом.

Элиминативная индукция – устранение альтернативных объяснений (т.е. случайностей, конфандеров и смещений – bias).

Суть *дедукции* и аналогии понятны [71]; первая применительно к критериям Хилла рассмотрена нами ранее [10].

В названной работе Vineis P., 1991 [71] автор, за малым исключением, избежал отнесения тех или иных критериев к конкретным типам разобранных им индукций и дедукций. По всей видимости потому, что многие критерии подпадают под несколько категорий, в зависимости от типа причинности [76], что и было рассмотрено здесь выше.

По признакам причинно-следственной связи, по типам доказательств, по источникам информации и по качеству доказательств

Авторы из USEPA в работе [77] и монографии [23] полагают, что различные критерии подпадают под указанные признаки или сочетают несколько таковых. Например, “Временная зависимость” подпадает под сочетание признаков: хотя хронологическая последовательность лучше всего анализируется в аспекте событий, поскольку причина всегда происходит до эффекта, для определения события требуется некоторая информация или предположения. “Сила ассоциации” относится к качеству доказательств, а “Эксперимент” служит источником информации. По подобным градациям в [23, 77] разнесены специфичные критерии причинности для экологии и экоэпидемиологии (о последних см. в [10]).

Сходные рассуждения приведены в [78]: только критерии “Сила связи”, “Временная зависимость” и “Биологический градиент” могут оцениваться на основе *данного конкретного исследования* без использования других источников. Прочие критерии Хилла связаны с дополнительной информацией.

Зависящие от статистической ассоциации и от биологической логики

Такие две категории упомянуты в Feinstein A.R., 1979a; 1979b [79, 80] для приведенных там пяти критериев (порядок как в оригинале): “Постоянство ассоциации”, “Сила ассоциации”, “Специфичность ассоциации”, “Временная зависимость”

и “Согласованность с известными фактами природной истории и биологии заболевания” (последний пункт как в Hill A.B., 1965 [11]). Три первых пункта “зависят от входных и выходных данных черного ящика, который называется “статистическими ассоциациями”, а “Временная зависимость” и “Согласованность” “зависят от “биологической логики” – “другого черного ящика” [80].

В главе по крайней мере двух изданий (2006 г. и 2018 г.) пособия по эпидемиологии рака S.N. Goodman и J.M. Samet также делили критерии Хилла на относящиеся к “прогнозируемым” (associational) – определяются экспериментально из изменений в вероятности возникновения патологии и “объяснительным” (explanatory) – основанным на предложенном механизме [81].

Вероятностные и генеративные критерии

В Thygesen L.C. et al. [82] критерии делятся согласно вероятностным закономерностям (probabilistic regularity) [83] и генеративному (порождающему) взгляду на причинность (generative view of causality) [84].

Вероятностные (“объект следует за причиной”): “Сила ассоциации”, “Постоянство ассоциации”, “Специфичность”, “Биологический градиент” и “Эксперимент”, причем последний понимается в широком смысле, не только как контрафактический (подробнее о типах эксперимента в критериях причинности см. в [7, 9]). Эти пункты “касаются наблюдаемых ассоциаций и, следовательно, связаны в первую очередь с вероятностной закономерностью причинности”.

Генеративные (“причина порождает объект”): “Биологическое правдоподобие”, “Согласованность с текущими фактами и теоретическими знаниями” и “Аналогия”. Данные пункты в первую очередь связаны с объяснительной силой сети взаимосвязанных ассоциаций и, следовательно, с генеративным взглядом на причинность [82]. Причинная связь здесь является частным случаем, выводимым из общих биологических законов, а не отношением, вызванным повторными наблюдениями (“...instance of more general laws of nature”) [85].

Критерий “Временная зависимость” не подпадает под указанные принципы, но играет центральную роль в определении направления эффекта [82].

Возникает сомнение относительно отнесения пункта “Эксперимент” в вероятностную категорию; возможно поэтому он просто выпущен при цитировании публикации [82] в пособии [86].

Вероятностные и механистические критерии

Смысл двух рубрик из Russo F., Williamson J., 2007 [87] похож на предыдущую градацию, однако полного совпадения входящих критериев не наблюдается.

Вероятностные соображения (probabilistic considerations): “Сила ассоциации”, “Постоянство ассоциации”, “Биологический градиент” и “Эксперимент” (во всех смыслах).

Механистические соображения (mechanistic considerations): “Временная зависимость” “Биологическое (в оригинале – теоретическое) правдоподобие”, “Согласованность с другими знаниями”, вновь “Эксперимент” (во всех смыслах) и “Аналогия”.

Особняком стоит “Специфичность”, которую в [87] вовсе отбрасывают как проблематичную (что не совсем обоснованно – см. в [8]).

Прямые, механистические и “параллельные” критерии

Эта градация предложена авторитетом в области ЕВМ J. Nowick и соавт. в 2009 г. [37] как развитие положений из [87] (далее порядок критериев соответствует оригиналу). Воспроизведены в [38].

Прямые (direct) доказательства: “Эксперимент”, “Сила связи” и “Временная зависимость” [37].

Механистические (mechanistic) доказательства: “Биологический градиент” и “Биологическое правдоподобие” [37] (подробно обсуждены в последующей работе Nowick J. et al., 2010 [88]).

“Параллельные” (parallel) доказательства: “Согласованность с текущими фактами и теоретическими знаниями”, “Постоянство ассоциации” и “Аналогия”. “Параллельные доказательства” – это которые подтверждают причинную гипотезу в соответствующих исследованиях, имеющих сходные результаты. Критерий “Специфичность” исключен вообще [37].

Подтверждающие причинность, отрицающие причинность и прочие

Эта градация для критериев в экологии и эпидемиологии была предложена в Fox G.A., 1991 [89], и затем развита в USEPA-1998 [61].

Строго подтверждающие причинность (strongly affirming causality): “Сила связи”, “Прогностическая эффективность (predictive performance)¹⁸”, “Биологический градиент” (stressor-response relationship) и “Постоянство ассоциации” [61, 89].

Последующие две градации обнаружены только в USEPA-1998 [61].

Обеспечивающие основу для отказа от причинности (providing a basis for rejecting causality): несоответствие критериям “Постоянство ассоциации” (Inconsistency in association), “Временная зависимость” (Temporal incompatibility) и “Согласованности с фактами” (Factual implausibility).

Другие уместные критерии (other relevant criteria): “Специфичность ассоциации” и “Теоретическое и биологическое правдоподобие”.

Логика отнесения к первой категории пункта “Прогностическая эффективность” не совсем понятна.

Концептуальные, конформационные, операционные и априорные

Эта философская рубрикация, причем для всех заболеваний, как инфекционных, так и хронических, предложена в Lower G.M., Kanarek M.S., 1983 [46].

Концептуальные (conceptual):

1. Причины заболевания. Возникновение заболевания связано с наличием или с воздействием реальных причинных агентов, переносчиков возбудителей или культурных особенностей, устанавливающих наличие опасности и риска. [“Ассоциация”¹⁹.]

2. Причины заболеваемости. Заболеваемость связана с культурно-средовыми детерминантами риска, определяющими степень воздействия и/или долю воздействия. [“Ассоциация”.]

3. Причины заболеваемости. Заболеваемость связана с культурно-средовыми детерминантами восприимчивости и хозяином, определяющими степень восприимчивости и/или восприимчивую фракцию.

4. Причины симптоматологии заболевания. Экспериментальное воздействие возбудителей болезни связано с возникновением симптоматики конкретного заболевания, включающего механизмы, относящиеся, если не идентичные, к тем, которые вовлечены в заболевание человека. [“Эксперимент”; на животных.]

Конформационные (conformational). Характеристики ассоциации на предмет причинности:

1. Постоянство (различные методологии, различная географическая локализация или ситуации воздействия).

2. Сила связи.

3. Согласованность с природной историей болезни и с биологическими механизмами.

4. Необходимость (Necessity): требует выполнения операционных критериев по предотвращению.

Операционные (operational). Все – в рамках “Эксперимент контрафактический” по предотвращению.

Априорные (a priori). То есть информация, которая уже имелаась до развития патологии (индуктивный подход). Входят в “Биологическое правдоподобие” и в экспериментальные подтверждения *in vitro*, в опытах на животных и на прочих модельных системах.

Ранги критериев по понятиям разных авторов и по частоте использования

Это, так сказать, “стихийный” показатель весомости критериев, основанный на никак не подкрепленных, но – мнениях отдельных исследователей и организаций. Как полагал D.L. Weed в 2000 г. [90], “выбор критериев – более вопрос общепринятой практики, чем теоретической строгости”. “Больше вопрос персонального предпочтения, чем тщательного исследования”²⁰.

Согласно M. Susser (1986) [41], критерии “нелегко (do not readily) формируют иерархию, хотя некоторые из них более сильны (forceful), чем другие”.

Кроме того, различные исследователи могут по-разному интерпретировать один и тот же критерий, как это было, например, с тремя определениями пункта “Постоянство ассоциации” при установлении связи между употреблением алкоголя и раком молочной железы [81].

Тем не менее попытки оценить важность тех или иных критериев имеются.

1. Научно-исследовательский совет США (National Research Council U.S.), в который входит АН США, в 1989 г. применительно к роли диеты при хронических патологиях использовал шесть критериев: “Силу связи”, “Постоянство ассоциации”, “Специфичность”, “Временную зависимость”, “Биологический градиент” и “Биологическое правдоподобие”. Всем критериям придавался равный вес, кроме – “Биологического правдоподобия”, получившего сниженную значимость, поскольку оно “зависит от субъективной интерпретации” [92] (цитировано по [93]).

2. В работе G.A. Fox от 1991 г. [89], по приложению принципов причинности в экологию и экоэпидемиологию, только три критерия Хилла называются как строго подтверждающие (affirm) каузальность: “Сила связи”, “Постоянство ассоциации”, “Прогностическая эффективность” (“Predictive performance”); это критерий M. Susser, а не Хилла; см. прим. 18) и “Биологический градиент”.

3. В пособии по эпидемиологии R.H. Fletcher и соавт., в 3-м издании от 1996 г. (перевод 1998 г.) [94] и в 4-м издании от 2005 г. [95], как основные названы критерии “Временная зависимость”,

“Сила связи”, “Постоянство ассоциации”, “Биологический градиент” и “Эксперимент контрафактический”. К дополнительным отнесены “Биологическое правдоподобие” и “Аналогия”.

4. В работах авторитета Weed D.L., 2000–2005 [50, 90, 96, 97] сказано, что в эпидемиологии рака наиболее важны для оценки причинной ассоциации “Сила связи”, “Постоянство ассоциации”, “Биологический градиент” и “Биологическое правдоподобие”.

5. В USEPA-2005 [45] как особо важные рассматриваются четыре критерия: “Сила связи”, “Постоянство ассоциации”, “Биологическое правдоподобие” и “Согласованность с текущими фактами и теоретическими знаниями (Coherence)”. Два других – “Временная зависимость” и “Биологический градиент” – также рассматриваются как актуальные, но упоминаются после.

6. В пособии ВОЗ по эпидемиологии Bonita R., 2006 [98] указано, что наибольший вес может быть придан “Биологическому правдоподобию”, “Постоянству ассоциации” и “Биологическому градиенту”.

7. В популярном пособии по эпидемиологии Gordis L., 2014 [39] к “главным” (“major”) отнесены критерии “Временная зависимость”, “Биологическое правдоподобие”, “Постоянство ассоциации” и “Альтернативные объяснения” (последнее – не критерий Хилла; подробно см. в [10]). К “прочим” (“other”) причислены “Биологический градиент”, “Сила связи” и “Эксперимент контрафактический”. Как указывал L. Gordis, данная градация²¹ была разработана еще в 1986 г. комиссией Минздрава США для оценки эффекта предродовых мероприятий, каковые критерии были опубликованы в Gordis L. et al., 1990 [99] и, затем, переиздавались в соответствующем пособии по эпидемиологии [39]. Этот достаточно странный список (“Сила связи” почему-то второстепенна) вошел (вероятно, из пособий L. Gordis) и в российский учебник Власов В.В., 2006 [44].

8. В пособии по эпидемиологии Bhopal R.S., 2016 [39] сказано: “Если бы испытание [здесь – trial в эпидемиологии] было бы свободно от всех ошибок и смещений (bias), то “Сила связи” и “Временная зависимость” отражали бы причинность”.

9. В главе о рисках от мобильной связи, включенной в монографию [100], в качестве приоритетных названы пять критериев Хилла: “Сила связи”, “Постоянство ассоциации”, “Временная зависимость”, “Биологический градиент” и “Согласованность с текущими фактами и теоретическими знаниями (Coherence)”. Вероятно, в последний пункт входит и “Биологическое правдоподобие”, но мы на этом акцентироваться не станем. (Оказалось, что применительно к эффек-

Таблица 1. Частота использования/упоминания девяти критериев Хилла в обзорах и работах
Table 1. Frequency of use/mention of the nine Hill's criteria in reviews and works

Исследование	Strength	Consistency	Specificity	Temporality	Biological gradient	Biological plausibility	Coherence	Experiment	Analogy
Weed D.L., Gorelic L.S., 1996 [43]. 14 обзоров от 1985–1994 гг.	13 (93%) 2 место	14 (100%) 1 место	5 (36%) 5 место	4 (29%) 6 место	9 (64%) 3 место	8 (57%) 4 место	2 (14%) 8 место	3 (21%) 7 место	1 (7%) 9 место
Настоящее исследование. 70 работ от 2013–2020 гг.	64 (91%) 2 место	64 (91%) 2 место	57 (81%) 4 место	65 (93%) 1 место	64 (91%) 2 место	59 (84%) 3 место	52 (74%) 5 место	59 (81%) 4 место	50 (71%) 6 место

там мобильных телефонов указанные пять критериев не выполняются даже близко [100].)

Это все, что удалось выявить по теме приоритетов для причинных принципов среди сотен работ. На основе цитированных девяти источников нами были присвоены ранги для отдельных критериев. Те, которые выше упоминались как основополагающие, получали 1 балл. Называвшиеся вторичными – 0.5 балла, а не упоминавшиеся вообще – 0 баллов. Порядок для суммы баллов получился следующий:

- “Постоянство ассоциации” – 8 баллов;
- “Сила связи” – 7.5 балла;
- “Временная зависимость” – 5.5 балла;
- “Биологический градиент” – 5 баллов;
- “Биологическое правдоподобие” – 5 баллов;
- “Согласованность (Coherence)” – 2 балла;
- “Эксперимент” – 1.5 балла;
- “Специфичность” – 1 балл;
- “Аналогия” – 0.5 балла.

Но это были как бы попутные оценки; остаются еще два специальных исследования – популярное, хотя и старое Weed D.L., Gorelic L.S., 1996 [43] (цитируется даже в последних пособиях по эпидемиологии [18, 40, 81]), и уже не раз приводившееся наше, которое ранее охватывало период 2013 г. – начало 2019 г. и основывалось на 35 публикациях [6, 8, 9], а в настоящей части 2 Сообщения 4 расширено до 2020 г. и до 70 работ. Поиск источников в нашем случае осуществлялся через PubMed на сочетание [“Hill criteria” & (Год)]. Цель обоих исследований заключалась в определении частоты использования/упоминания того или иного критерия в обзорах [43] или в конкретных работах, где пункты Хилла служат для доказательности (наше исследование). Результаты обоих изысканий представлены в табл. 1.

Исследование 1996 г. [43] и наше весьма расходятся в результатах. Конечно, какие бы оценки и ранги ни присваивались, очевидно, что критерий “Временная зависимость” априори должен сто-

ять на первом месте [5, 8]. Но в реальности приведенная выше балльная оценка, вкуче с работой [43] и нашими данными (табл. 1), показывают, что на 1–2-м месте находится сугубо индуктивный принцип – “Постоянство ассоциации”, а дедуктивный “Эксперимент” – во всех случаях на третьем месте с конца.

“Сила связи”, также во всех случаях, занимает второе место (или с отрывом, или, как в нашем исследовании, совместно с “Постоянством ассоциации” и “Биологическим градиентом”). Исходя, таким образом, из просто мнений и использований критериев Хилла, можно выстроить следующую ранговую зависимость для самых необходимых причинных пунктов:

- “Временная зависимость”;
- “Постоянство ассоциации”;
- “Сила связи”;
- “Биологический градиент”;
- “Биологическое правдоподобие” (вкуче с “Согласованностью”).

Отметим, что необходимость последнего пункта подкрепляет важность теоретических и экспериментальных исследований любого уровня, включая радиобиологические.

Странным образом, как сказано, остался на последних местах называвшийся А.В. Hill “самой сильной поддержкой гипотезы причинности” [7, 9, 11] “Эксперимент контрафактический”, т.е. “природный эксперимент” (“semi-experimental”): наблюдение за эффектом, когда его гипотетическая причина по естественным причинам (точнее, не зависящим от исследователя) ослаблена или устранена. А ведь контрафактический (“манипулятивный”) подход рассматривается как одна из основных методологий доказательства причинности [12–14, 19, 21–27] (см. также далее). По-видимому, дело в том, что уповать на вероятность “природного” эксперимента, который сам не поставишь, все равно что играть в лотерею. Хотя большинство исследователей явно не понима-

ют разницу между контрафактическим и обычным экспериментом [7, 9].

В работе Holman C.D., 2001 [101] по опросу определяли, какие положения или критерии суждения о причинности являются приоритетными применительно к связи между факторами окружающей среды и патологиями. Имелось 12 примеров из практики; контингент составил 159 эпидемиологов Австралии.

Факторы, наиболее влияющие на выводы о причинной связи, были следующие (OR (отношение шансов) рассчитаны в [101] для суждений эпидемиологов относительно значимости тех или иных приведенных диапазонов RR (относительный риск) и *p*, как если бы таковые им встретились в публикациях):

- значимость критерия “Ассоциация” – статистическая значимость: отношение шансов – OR = 4.5, для $0.001 < p < 0.05$ и OR = 7.2, для $p < 0.001$ против $p \geq 0.05$;
- значимость критерия “Опровержение альтернативных объяснений”: OR = 8.1 для неизвестного конфаундера против его отсутствия;
- значимость критерия “Сила связи”: OR = 2.0, для $1.5 < RR \leq 2.0$ и OR = 3.6, для $RR > 2.0$; все против $RR \leq 1.5$;
- значимость критериев “Биологическое правдоподобие и “Согласованность”: дополняющая информация о биологической, фактической и теоретической согласованности.

Другими значимыми переменными были количество исследований (“Постоянство ассоциации”), зависимость “доза–эффект” (“Биологический градиент”) и тип (дизайн) исследования. Видим, в основном, все те же приоритетные критерии, что и выше.

Отдельные авторы попытались использовать ранговый подход для непосредственного взвешивания критериев с последующим практическим применением.

Взвешивание критериев Хилла

Про принципиальную возможность взвешивания отдельных критериев говорилось рядом авторов, но без указания на то, как это следует делать [49, 66]. M. Susser привел в 1986 г. [41] свои соображения по значимости конкретных критериев, но в его исследовании даны только степени поддержки/отклонения/нейтралитета для исходной дедуктивной гипотезы при выполнении каждого критерия (табл. 2).

Похожая градация была предложена ранее, в 1981 г., McMaster University Health Sciences Centre (Канада) [102]. Заметим, что в этом списке “Биологическое правдоподобие” и “Согласованность” заменены на “Биологический смысл” и “Эпиде-

Таблица 2. Полуколичественная оценка значимости при поддержке причинной гипотезы. По Susser M., 1986 [41] **Table 2.** Semi-quantitative assessment of significance in support of the causal hypothesis. According to Susser M., 1986 [41]

Критерий и его варианты	Эффект на гипотезу		
	Поддержка	Отклонение	Неопределенно
Временная зависимость (“Time order”)			
Выполняется			+
Не выполняется		++	
Неопределенно			+
“Сила связи” (RR или корреляция согласно [41]*)			
Очень сильная	++		
Сильная	+		
Слабая			+
Специфичность			
<i>Специфичность эффекта</i>			
Высокая	+/-		
Низкая			+
<i>Специфичность причины</i>			
Высокая	++		
Низкая			+
Постоянство ассоциации (“Consistency of replication” согласно [41])			
Высокая позитивная	++		
Высокая негативная		++	
Низкая			+
Согласованность с текущими фактами и теоретическими знаниями (“Coherence”)			
<i>С теорией</i>			
Правдоподобно	+		
Неправдоподобно		+	
<i>С биологией</i>			
Согласуется	+		
Не согласуется		+	
<i>С фактами</i>			
Совместимо	+		
Несовместимо		++	
<i>Со статистикой</i>			
Дозовая зависимость монотонная	++		
Дозовая зависимость другая			+

* В эпидемиологических исследованиях корректным показателем “Силы связи” служит RR, а не коэффициент корреляции (подробно см. в [3, 4, 8]).

Таблица 3. Оценка значимости дизайнов исследований и каузальных критериев для доказательства причинности связи. По [102] (1981)

Table 3. Assessment of the significance of research designs and causal criteria for proving causality. According to [102] (1981)

Тест, критерий	Эффект на причинность		
	поддержка причинности	нейтрально или не поддерживает причинность	отклоняет причинность
Опыты на людях	++++ (4 плюса)	--- (3 минуса)	---- (4 минуса)
Сила ассоциации			
RCT	++++ (4 плюса)	--- (3 минуса)	---- (4 минуса)
Когортное исследование	+++ (3 плюса)	-- (2 минуса)	--- (3 минуса)
“Случай – контроль”	+ (1 плюс)	0 (нуль)	- (1 минус)
Постоянство ассоциации	+++ (3 плюса)	-- (2 минуса)	--- (3 минуса)
Временная зависимость	++ (2 плюса)	-- (2 минуса)	---- (4 минуса)
Биологический градиент	++ (2 плюса)	- (1 минус)	-- (2 минуса)
Согласованность (“Coherence”); в [102] “Epidemiologic sense”	++ (2 плюса)	- (1 минус)	-- (2 минуса)
Биологическое правдоподобие ; в [102] “Biologic sense”	+ (1 плюс)	0 (нуль)	- (1 минус)
Специфичность	+ (1 плюс)	0 (нуль)	- (1 минус)
Аналогия	+ (1 плюс)	0 (нуль)	0 (нуль)

миологический смысл”, как было предложено в 1978 г. D.L. Sackett [103] (подробнее см. в [10]). Канадская иерархия выходит за рамки только критериев причинности, охватывая также и дизайны исследований, причем как в области эпидемиологии, так и ЕВМ (табл. 3).

Представленные в табл. 2 и 3 градации могут быть спорны, особенно из Susser M., 1986 [41]. Но, располагая множеством источников по иерархии доказательств причинности, мы, кроме публикации 1981 г. [102], более не встретили схемы, охватывающей одновременно и дизайны исследований, и критерии Хилла.

В работах из области экологии и экоэпидемиологии [104–106] приведена система оценки веса свидетельств (WoE), в которой используется скоринг (“score”, бальное ранжирование) различных типов доказательств. Индексирование присваивает веса различным типам доказательств, чтобы объединить их в единое значение, которое определит результат анализа. Отмечается, что при определении весов “Лучшее профессиональное суждение” (“Best Professional Judgment” – BPJ) основывается на критериях “Постоянство ассоциации”, “Специфичность” и “Сила ассоциации” [105].

В USEPA-2000 [107] приведена качественная оценка значимости (также “score”) для отдельных критериев в области экоэпидемиологии (подроб-

нее о таких критериях см. в [10]). Оценка основана на присваивании соответствующего количества символов “+”, “–” или “0”, связанных с силой доказательства для каждого пункта, оцененного для каждой вероятной причины [104, 107]. Указано, что не рекомендуется складывать баллы по каждому критерию, поскольку сложение баллов ошибочно означает, что все соображения одинаково важны. Последнее возможно только в том случае, если для всех вероятных причин у USEPA имеются одинаковые типы доказательств [107].

Разобраться в этом нелегко, тем более, что в последующих документах USEPA (см. в [10]) подобное качественное взвешивание, насколько нам известно, уже не встречается.

Количественное взвешивание отдельных критериев Хилла разработано в экотоксикологии, в рамках международной программы ВОЗ по химическим токсикантам в окружающей среде (The International Programme on Chemical Safety – IPCS), которая выполняется с 1980 г. (см. в [10]). В Becker R.A. et al., 2017 [108] приведены следующие удельные веса критериев (порядок как в оригинале; доли заменены нами на проценты):

- “Биологическое правдоподобие” – только как “ознакомительное” (narrative);
- “Эксперимент контрафактический” (“Essentiality of Key Event”²²) – 40%;

- “Биологический градиент” (“Dose and Incidence Concordance”) – 20%;
- “Временная зависимость” (“Temporal Concordance”) – 20%;
- “Постоянство ассоциации” – 10%;
- “Аналогия” – 10%.

Как видим, набор критериев, равно как и ранги значимости, для экотоксикологии специфичен: на первом месте упомянутый выше “природный эксперимент”, которого еще надо дожидаться. Не рассматривающийся серьезно в эпидемиологии критерий “Аналогия” [9] получает здесь вполне важный статус.

Несколько другие ранги и веса в работе Collier Z.A. et al., 2016 [109], также по оценке WoE в экотоксикологии:

- “Биологическое правдоподобие” (“Biological Concordance”) – 24%;
- “Эксперимент контрафактический” (“Essentiality of Key Event”) – 29%;
- “Временная зависимость” + “Биологический градиент” + специальный показатель для Key Events (суммарно: критерий “Concordance of Empirical Observation among Key Event”) – 22%;
- “Постоянство ассоциации” – 13%;
- “Аналогия” – 12%.

Видно несоответствие между оценками в [108] и [109] значимости критериев “Биологическое правдоподобие”, “Временная зависимость” и “Биологический градиент”. В целом вновь разобраться в этом и выбрать нечто унифицированное затруднительно. Но в который раз есть простор выбора для многих конъюнктур.

Наиболее разработана методика количественного взвешивания критериев Хилла, по-видимому, в не получившей никакого развития работе нидерландских авторов Swaen G., van Amelsvoort L., 2009 [55]²³. Позднее, в обзоре по методикам оценки WoE, подход из [55] был назван “Weighted Bradford Hill” [28].

Суть подхода из [55] следующая.

1. Вероятность того, что критерий будет соблюден, должна рассчитываться на основе имеющихся эпидемиологических и иных данных. В большинстве случаев имеющиеся свидетельства в поддержку не будут иметь вес в 100%. Например, данные о постоянстве ассоциации могут быть оценены как 75% в ситуации, когда три из четырех исследований сообщают о связи, а четвертое – нет. Критерий “Биологическое правдоподобие” может быть оценен с вероятностью 60%, если только одно исследование на животных продемонстрировало сходные результаты. Этот критерий должен рассматриваться с позиции наиболее полного набора данных на животных, экспериментов *in vitro* и информации о возмож-

ных механизмах. Критерий “Аналогия” оценивается как выполненный на 0%, если хорошо спланированные исследования аналогичного соединения дали отрицательные результаты.

2. Взвешивающий метод, основанный на критериях Хилла, требует двух параметров.

а) Вероятности того, что каждый из девяти критериев является истинным для данной эпидемиологической (и другой токсикологической) базы данных. Это – предмет экспертного суждения.

б) Расчета относительного веса для вклада каждого критерия в общий причинный вывод. Это – предмет эмпирического подхода.

Авторы [55] следующим образом подошли к каждому критерию.

“Силе связи” придавалась вероятность в зависимости от величины RR: 60% при 1–2; 80% – при 2–5 и 95% – свыше 5. Если RR из нескольких исследований попадали в разные категории оценки, то выбиралась представляющая наиболее надежную и последовательную оценку RR. Но формальный мета-анализ не проводился.

Постоянство ассоциации определялось по числу исследований: 0% для одного, 75%, если для 3 из 4 работ результаты оказывались последовательными, 95% – когда 4 и более публикаций показали одинаковый эффект. Если данные работ были противоречивыми, но общий эффект являлся положительным, то вероятность оценивалась как 60%.

“Специфичность” получала высокий балл в случае единственного результата действия агента. Например, для мелкоклеточного рака легкого при воздействии эфира 1,4-бис(хлорметил)-бензола вероятность оценивалась как 95%. В случае солей кадмия, где в некоторых исследованиях было выявлено повышение частоты рака предстательной железы, а в других – рака легкого и молочной железы, вероятность специфичности была оценена в 40%. Для некоторых агентов эпидемиологические доказательства состояли только из сообщений о случаях заболевания, но все они указывали на сходный тип рака. В таком случае вероятность оценивалась в 60%.

“Временная зависимость” при наличии эпидемиологических данных почти всегда соблюдалась и имела вероятность 100%.

“Биологическому градиенту” придавались 90%, если одно или несколько исследований четко сообщали о положительной зависимости “доза – эффект”. Когда такого ответа получено не было, но исследование проводилось, то вероятность оценивали в 30%. Если одна работа показала отчетливый ответ на дозу, а другая – нет, то назначали 50%. Если для агента никогда не сообщалось о зависимости “доза – эффект”, или исследова-

ния не продемонстрировали зависимость, то вероятности придавали 0%.

“Биологическое правдоподобие” основывалось на наличии экспериментов на животных или данных о механизме. При наличии существенных данных для животных критерию придавали 90%. Если имелось только одно такое исследование, оценка составляла 80%. Если были соображения, почему канцерогенный эффект у животных не может быть перенесен на человека, то вероятность оценивалась в 60%. Если единственное доступное свидетельство состояло, скажем, в подавлении репарации ДНК, то вероятность оценивали в 50%. Если было показано, что иммуносупрессивный агент влияет на иммунную систему, то критерий правдоподобия оценивался с вероятностью 40%.

“Согласованность” с текущими фактами и теоретическими знаниями”. Ранние эффекты или другие нарушения в органе-мишени способствуют критерию согласованности. Например, данные, что бензол вызывает повреждение костного мозга, указывают на то, что критерий согласованности был выполнен с вероятностью 80%.

“Эксперимент контрафактический”. Согласно [55], только для немногих агентов имеются данные о снижении избыточной частоты рака после прекращения воздействия. Основным примером, с многочисленными данными, является прекращение курения (подробнее см. в наших обзорах [2, 7, 9]). В подобном случае вероятность оценивалась в 95%.

Аналогия. Оценка вероятности основывалась здесь на канцерогенном потенциале иных подобных агентов, например, других алкилирующих соединений. Вероятность в этом случае критерия аналогии для интересующего алкилирующего агента оценивалась в 60%. Для иммуносупрессивного средства эта вероятность была определена в 20%.

Вес каждого из критериев определялся с помощью дискриминантного анализа.

Конечно, значительная субъективность приведенных оценок вероятностей бросается в глаза (что признают сами авторы [55]), как и, в целом, для байесовских подходов [13, 17, 52, 110, 111]. Но практическое применение показало приемлемость методики. Авторы [55] оценивали по указанным весовым критериям Хилла вероятность канцерогенных эффектов для 159 агентов, выбранных по списку Международного агентства по изучению рака (МАИР; IARC) в категориях 1 и 2А (т.е. с наибольшими рисками). Агенты в категориях 2В, 3 и 4 не включались, поскольку для большинства из них просто нет эпидемиологических данных.

Как сказано в [55], “в целом прогностическая модель сработала хорошо”. Частота согласованных (точных) прогнозов составила 82% (130 со-

единений из исследованных 159). Модель верно классифицировала 89% (85 из 96) канцерогенов IARC категории 1 и 71% (45 из 63) канцерогенов категории 2А. Три критерия Хилла внесли наибольший вклад:

- “Сила связи”;
- “Постоянство ассоциации”;
- “Эксперимент контрафактический”.

Эти три критерия в совокупности объясняли 41% общей дисперсии в базе данных.

Авторы [55] указывают, однако, что разработанная модель может быть применена для соединений и воздействий типа категорий 1 и 2А по МАИР, поскольку для остальных категорий, как сказано, мало эпидемиологических данных и, потому, наиболее важные эпидемиологические критерии априори внесут меньший вес. Другим ограничением является то, что модель все еще требует большой степени экспертного суждения [55].

Мы можем добавить также возможность “круговой логики” [112, 113] при проверке методологии из [55] на канцерогенных соединениях по категориям МАИР. Дело в том, что сам МАИР часто использует для подтверждения выводов об опасности соединений именно критерии Хилла [114, 115].

Однако, как и в предыдущем подразделе, где рассмотренные ранги критериев сформировались “стихийно” (только на основе мнений), модель из [55], предусматривающая более объективную оценку, базирующуюся на эпидемиологических данных, вывела на вершину все те же критерии. При исключении, понятно, заведомо 100% по любым оценкам “Временной зависимости”, на первых местах стоят “Сила связи” и “Постоянство ассоциации”. Индуктивные принципы [8, 10].

Представленные здесь материалы по придумыванию процентных вероятностей-весов, на первый взгляд, могут выглядеть несерьезными и игрушечными для экспериментатора, занимающегося лабораторными изысканиями на гипотетико-дедуктивной основе, с рукотворным формированием твердых “контролей” — “опытов”. Но следует помнить, что большинство исследований в области эпидемиологии, общественного здравоохранения, радиационной безопасности и т.д. являются *обсервационными*, поскольку нельзя провести не только рандомизированные контролируемые испытания (RCT), но и просто какие-либо эксперименты на людях, даже на добровольцах [22, 27, 116–118] (и мн. др.). В значительной степени та же ситуация — в экологии, экоэпидемиологии и экотоксикологии [23, 45, 61, 104–109]. Однако ответственные организации, принимающие решения, обязаны опираться на некие количественные или полуколичественные оценки рисков, по крайней мере в рамках “предупре-

дительного принципа” [1]. Отсюда и обилие стандартизованных способов и методик WoE, “Веса свидетельств”, спектр которых весьма широк и насчитывает массу методологий от множества международных организаций, комиссий, комитетов, программ и т.п.

В исследовании Martin P. et al., 2018 [28] были проанализированы соответствующие документы 63 национальных и международных агентств и 116 иных руководств-публикаций, устанавливающих WoE. Всего нашлось 24 методологии, включая классические [11], модифицированные [10, 106–109] и взвешенные [55] критерии Хилла. Забегая вперед, к части 3 настоящего сообщения, можно отметить, что, уже по нашему исследованию, более половины из оставшегося 21 подхода, перечисленного в [28], сводятся ко все тем же критериям Хилла, а среди незатронутых значительная часть – это просто обзоры, мета-анализы и экспертные заключения.

Таким образом, реальность не совсем совпадает с тем, что представляется лабораторному экспериментатору серьезным или несерьезным в плане доказательности. Необходимость практических решений делает даже выглядящие, порой, субъективными и искусственными подходы вполне востребованными, а потому – серьезными. В отличие от иных обсервационных дисциплин, являющихся реконструктивными и ретроспективными (история, археология, палеонтология и пр. [119]), истинность кажущегося субъективным “Веса свидетельств” (WoE) в эпидемиологии подтверждается проспективно – на будущих наблюдениях и реалиях практики.

КРИТИКА ПРИЧИННЫХ КРИТЕРИЕВ

Материалы в конце предыдущего подраздела, являющиеся, как сказано, некоторым забеганием вперед, показалось необходимым привести в том числе потому, что у ознакомившегося с настоящим критическим разделом может создаться впечатление, будто “современная эпидемиология давно отрунула всяческие критерии причинности”, что это “прошлый XX век”, что “есть много иных подходов” и что “ведущие мировые авторитеты давно развенчали...”, так сказать. При желании, тенденциозно опустив материалы, изложенные нами в [1–10], и опираясь только на представленное в данном разделе, можно похоронить весь индуктивный подход и, вкупе с ним, все критерии причинности от A.V. Hill, M. Susser, A.S. Evans и других авторов [1–11, 30–33, 41–43, 72, 74, 89, 92, 105–108, 114, 115] в эпидемиологии и иных обсервационных дисциплинах. Мы предупреждаем, что подобный конъюнктурный подход, а особенно – в области радиационных дисциплин (где контролируемые эксперименты на людях в большинстве случаев невозможны [118]),

потребуется осознания, что же у авторов имеется взамен для доказательности. Как будет видно в следующем разделе, это “взамен”, на наш взгляд, настолько сложно, неоднозначно трактуемо и, порой, трудно применимо на практике, что желающим оказаться “на пике современной эпидемиологической моды” либо просто затенить недостатки своего исследования, необходимо хорошо подумать, прежде чем, получив как бы “последнее знание”, откидывать методологии, основанные на причинных критериях.

Множество международных организаций и комитетов, как сказано, не смогли пойти по упомянутому “модному пути” [28] (это мы еще рассмотрим в части 3 Сообщения 4). Поскольку основа их деятельности – практика и принятие решений, а не только теоретические изыскания вокруг философии науки, эпидемиологии, математики и статистики. Позволим себе, отойдя от академического стиля, выразить здесь собственное мнение, которое, конечно, субъективно, и кто-то может обвинить нас в прагматизме²⁴. Много из приведенного ниже материала от теоретиков науки и, конкретно, эпидемиологии, напоминает притчу о продвинутых философах, бытовые и жизненные дела за которых выполняют их менее продвинутые и менее умные домочадцы. Причем, как мы увидим в соответствующем подразделе, “посвященные философы” полагают, что *это так и должно быть* [48, 110, 120, 121].

Согласно философии, в естественнонаучных дисциплинах невозможно истинное знание о причинности, а только – вероятностное

“Если мы просим доказательств в медицине или в любой другой эмпирической науке, то мы просим того, чего не существует”.

J. Cornfield, 1954 [122]²⁵

Истинное (абсолютное) знание о причинно-следственной связи получить невозможно ни в обсервационных, ни в экспериментальных дисциплинах. Этот вопрос рассматривался нами ранее [1, 3, 5, 7–9]. Никакой тщательно спланированный опыт, даже в лаборатории, здесь не поможет, поскольку нет возможности сформировать настоящий “контрафактический идеал”, и он заменяется суррогатным “контрафактическим контрастом”, т.е. реальным контролем²⁶.

Еще Д. Юм в XVIII в. указывал, что строгое доказательство в эмпирической науке – невозможно [125]; позже это повторил и К. Поппер [126].

Рассуждения о неабсолютности причинно-следственного знания в естественных науках и в эпидемиологии в частности можно найти во многих источниках [47, 49, 59, 121, 123, 127, 128],

включая работу Hill A.V., 1965 [11], соответствующая цитата из которой, отображенная аутентично в массе западных публикаций и пособий [18, 25, 33, 35, 39, 40, 47, 54, 57, 59, 89, 129–131], приводилась нами ранее [7, 9]²⁷.

Отсутствие абсолютного знания, однако, приводит к мировоззренческим издержкам. Как отмечал философ науки I. Lakatos [132]:

“Ныне очень немногие философы или исследователи все еще полагают, что научное знание является доказанным или может быть таковым. Но мало кто понимает, что иначе вся классическая структура интеллектуальных ценностей рушится, и она должна быть заменена: нельзя, как это делают некоторые логические эмпирики, просто разбавить идеал доказанной истины до идеала “вероятной истины”, или, как это делают некоторые социологи знания, свести к “истине путем [изменения] консенсуса”²⁸.

Высказывание I. Lakatos послужило S.F. Lanes в 1988 г. основанием для пессимистического заключения о возможностях эпидемиологии вообще и причинных критериев – в частности [123].

Знание получить, таким образом, невозможно, но если очень надо – то, как показывает бытие, вполне можно. На данный диссонанс обратили внимание и иные авторы [127] (курсив мой. – А.К.):

“С одной стороны, много усилий потрачено, чтобы тщательно разработать методы, демонстрирующие причинные отношения, в то время как, с другой стороны, нам говорят – довольно убедительно – что мы не можем в конечном счете определять причинные отношения, или что мы должны воздержаться от стремления установить причинные отношения, потому что они должны быть поняты как *отличные* от законополагающих (nomological) или *вероятностных* связей”.

Вероятностная и детерминированная природа причинности. Суть приведенной цитаты в понятии “вероятностных”. В обзоре [1] нами рассматривались пять концепций причинности [35, 133], от простейшего толкового (“по действию”) до сложных комплексных (необходимые и достаточные причины, компонентные причины), а также контрафактические причины и вероятностные причины²⁹. Последнее понятие имеется уже у Д. Юма [125], вполне применимо также к компонентным причинам [1], и наиболее подходит для эпидемиологии, оперирующей частотой эффекта на популяцию, но не индивидуальными рисками [9, 10]. Если в качестве показателя взять недетерминированные последствия воздействий (например, радиации), т.е. канцерогенез и трансгенерационные эффекты [65, 118], то очевидно, что только вероятностная причинность и окажется актуальной.

Современное понятие о причинности в медицине [134] и эпидемиологии [135] в целом предусматривает именно вероятностную позицию.

Между тем, ряд авторов стоят на концепции детерминированной каузальности в эпидемиологии. По их мнению, хотя причинные связи и кажутся вероятностными на практике, это может быть обусловлено незнанием полного набора компонентов, составляющих сложную достаточную причину патологии [48, 136–139]. В двух первых изданиях “Modern Epidemiology” (1986 г. [137] и 1998 г. [138]), которые нам недоступны, K.J. Rothman (во втором случае с S. Greenland), согласно цитирующим источникам, выступают за детерминистскую концепцию причинности для своего многокомпонентного каузального “пирога” [13, 52, 53, 110, 137, 138] (подробно про последний см. в [1] и ниже).

В работе 2004 г. [140] приводятся слова K.J. Rothman из первого издания названного пособия [137], что 10% вероятность рака легкого у активных курильщиков обычно оценивается, как если бы результат, помимо курения, был чисто случайным. Но конструктивнее рассматривать распределение равных рисков как отражение не более чем *нашего незнания детерминант* рака легкого, которые взаимодействуют с сигаретным дымом. У K.J. Rothman выходит так, что подобные детерминанты *когда-нибудь можно будет узнать*: по мере расширения знаний, оценки риска, присвоенные людям, будут приближаться *к нулю или единице* [137] (цитировано по [140]).

В работе 2011 г. [141] представлены сходные рассуждения, но уже из второго издания “Modern Epidemiology”, от 1998 г. [138]. В [141] сказано следующее (курсив мой. – А.К.):

“*Удивительно*, что оценки каузальности в эпидемиологической литературе не полностью ограничиваются (embraced) вероятностным подходом. Например, важная (influential) составляющая компонентной модели Rothman и Greenland является, в основном, детерминированной... Хотя данные авторы (Rothman K.J., Greenland S. 1998 [138]) признают, что причинно-следственные связи часто представляются на практике вероятностными, *это связано с незнанием полного набора компонент*, составляющих сложную, достаточную причину заболевания”.

В обзоре Davey Smith G., 2019 [142] подтверждается приверженность K.J. Rothman именно детерминистской концепции в его “пироге причинности” как в первом, там и во втором изданиях “Modern Epidemiology” [137, 138].

В имеющемся у нас третьем издании “Modern Epidemiology” от 2008 г. [13] рассуждения про курение более обтекаемы, но в главе из пособия по эпидемиологии от 2014 г. [143] обнаружены те же рассуждения K.J. Rothman и S. Greenland (курсив

мой. – А.К.): “Мы не можем измерить индивидуальные риски, и присвоение среднего значения каждому в категории отражает не что иное, как наше незнание детерминант рака легкого, которые взаимодействуют с сигаретным дымом”.

Таким образом, предлагаемые К.Ж. Rothman и S. Greenland оценки причинности именно детерминизм и подразумевают. В обзоре Davey Smith G., 2019 [142] попытки вывести детерминированную причинную атрибутивность каждого эффекта названы “фантазией”, поскольку игнорируются ограничения материального мира (“Мы потворствуем (encourage) мифопоэтике персонализированной (personalized) медицины; эпидемиологи, конечно, должны относиться к этому подозрительно”)³⁰.

Желание доискаться до всех конкретных причин эффекта выглядит странно и с позиции радиационной эпидемиологии, тем более на современном этапе превалирования стохастических последствий облучения [65, 115, 118].

Столь углубленная информация по концепциям К.Ж. Rothman и S. Greenland обусловлена, во-первых, их отрицанием критериев причинности [13, 52, 110, 137, 138]³¹ и, во-вторых, декларируемым высоким авторитетом по крайней мере первого автора в современной эпидемиологии, отражающимся в весомых источниках последних лет:

“Основной причинной моделью, используемой сегодня эпидемиологами, являются “пирогги” К.Ж. Rothman” (2006) [19]³².

“В то время как критерии Хилла дают некоторую полезную эвристику для определения причинности, нынешняя система эпидемиологии в значительной степени определяется мышлением К.Ж. Rothman” (2016) [86]³³.

“Монография К.Ж. Rothman 1986 г. “Modern Epidemiology” [137] стала переломным моментом (watershed) в этой дисциплине” (2019) [142]³⁴.

Как было разобрано выше, К.Ж. Rothman исповедует причинный детерминизм. Между тем, А.В. Hill с самого начала указывал, что его “положения” (“viewpoints”) – это не строгие правила-критерии, а некие принципы, руководства, чтобы оценить степень вероятности того, что ассоциация причинна [11]. Кроме “Временной зависимости” [5, 8], детерминизма критерии не предусматривают.

Никакой индукции: только гипотетико-дедуктивный метод К. Роррег. Никаких причинных критериев

Ряд причинных критериев прямо индуктивны, по крайней мере “Ассоциация” (критерий М. Susser [10]), “Сила связи”, “Постоянство ассоциации”, “Специфичность” [71] и “Биологический градиент” (который иногда рассматривается

в рамках “Силы связи” [3, 27, 31, 59]). Хотя, как было видно выше, согласно Weed D.L., 1988 [76] большинство критериев, кроме “механистических” [37, 38, 87] (“Правдоподобие”, “Согласованность”, “Аналогия”), могут подвергаться и дедуктивному подходу в зависимости от типа причины.

В целом причинные критерии рассматриваются как индуктивно-ориентированные [13, 52, 110, 144].

Но для ряда авторов даже в эпидемиологии единственный научный подход к доказательности – это гипотетико-дедуктивный метод К. Роррег [76, 91, 121, 123, 147–152].

«“Эпидемиологические критерии причинности” были интеллектуальным позором, а уровень аргументации... иногда был более достоин литературных критиков, чем ученых» (1993 г., философы) [153].

“Критерии “установления причинности” и другие методы субъективного вывода не служат полезной цели” (1988) [121].

“Индукции нет и в эпидемиологии” (1988) [150].

“Таким образом, авторитетные специалисты в области эпидемиологии отрицают причинную убедительность критериев, которые не имеют силы в качестве первичных выводных детерминант причинности, а просто остаются атрибутами *ex post facto* причинных заключений, которые ранее были подтверждены экспериментально” (2004) [140].

«Обращение... к непроверенным “критериям” великого светила из прошлого напоминает “научные” методы Темных веков (Dark Ages)» (2004) [33].

“Очевидно, что стандарты эпидемиологических данных, предлагаемые Хиллом, обременены оговорками и исключениями” (2005) [52, 110].

“[К.Ж. Rothman] резко критикует критерии Хилла, и заключает, что контрольные списки придают обманчивый и бессмысленный авторитет изначально несовершенному и творческому процессу предположений и опровержений” (2012) [78].

“Одним из важных ограничений классического взгляда на причинно-следственную связь заболелания, вытекающего из критериев Хилла, было отсутствие формальной основы для оценки причинных гипотез” (2013) [22].

“Большинство критериев Хилла имеют спорную и весьма условную применимость” (2013) [154].

“Критерии Хилла не могут рассматриваться как необходимые или достаточные условия для превращения ассоциации в причинную ассоциацию” (National Research Council USA, 2014) [155] (цитировано по [156])³⁵.

«Этот список был общепринятым стандартом в эпидемиологических рассуждениях и использовался почти как “доказательство” до тех пор, пока К.Ж. Rothman (1986 [137]) не только не раскрити-

ковал все положения как неубедительные, но и заменил их детерминистской теорией причинности связей для эпидемиологических исследований» (2016) [86].

И другие цитаты в этом роде; см. также в [9].

Однако в обзоре [9] нами рассматривались проблемы и практическая виртуальность “Эпидемиологии Поппера” (“Popperian Epidemiology”), которая живо обсуждалась и насаждалась как философами, так и отдельными эпидемиологами в 1970-х–1980-х гг. [76, 120, 121, 123, 144–152], а позднее как-то почти сошла на нет.

Некоторые авторы (M. Susser в 1986–1995 гг. [41, 42, 144–146], D.L. Weed в 1986–1988 гг. [76, 147, 148]; также упоминание в 2008 г. [149]) попытались “спасти” от непримиримых приверженцев К. Поппера критерии Хилла, добавим к ним дедуктивные пункты “Прогностическая эффективность” (“Predictive performance”; M. Susser), а также “Предсказуемость” и “Тестируемость” (“Predictability” и “Testability”; D.L. Weed). В [9] мы указывали, что применение подобных положений, расплывчатых и направленных как бы “в будущее”, проблематично, когда надо оценить истинность ассоциации для практических шагов, и не задним числом.

Тем не менее жизнь и практика эпидемиологии брали свое. В [9] нами подробно, с приведением весомых источников, рассматривались вопросы превалирования именно индуктивного подхода в классической эпидемиологии. И в нашей жизни. Какой бы “современной” [13, 137, 138] и “постсовременной” [142] эпидемиология не стала, пока не видно, по крайней мере в западных и российских тематических пособиях (нам доступно порядка 70; см. отчасти в [3–10]), каких-либо сдвигов в сторону не только “Popperian Epidemiology”, но даже в сторону широкого использования иных подходов определения причинности (рассмотрены ниже). По крайней мере таково наше впечатление. И это — не говоря уже о документах международных организаций, имеющих дело с лучевым фактором [65, 118, 157].

Критерии причинности — не для науки, а только для практики. Оставить решения политикам и соответствующим лицам

“Эпидемиологи, похоже, путают практические результаты эпидемиологических исследований на популяционном уровне с метафизическими взглядами на реальность причинности патологий на индивидуальном уровне”.

De Vreese L., 2009 [158]³⁶

V.G. Charlton, взгляды которого на критерий “Биологическое правдоподобие” мы рассматривали ранее [6, 9], в 1996 г., категорически отрицая критерии Хилла, указывал, что они полезны только для того, чтобы давать рекомендации в области общественного здравоохранения (“политические решения”), и что их использование делает эпидемиологию менее научной, более мультидисциплинарной, производящей ненадежные и не подлежащие обобщению результаты [48].

В 1988 г., после симпозиума по “Popperian Epidemiology” (см. в [9]), D.L. Weed заявил следующее: “...я согласен с тем, что правила медицинской технологии (medical technology), в рамках которой лежит общественное здравоохранение, и правила медицинской науки, хотя и сходны, но не идентичны. Правила технологии подчеркивают подтверждение; правила науки подчеркивают опровержение” [148].

Это квинтэссенция: “*практика подчеркивает подтверждение, наука подчеркивает опровержение*”.

Образно, хотя и утопично. Собрался, скажем, кто-то подтвердить, что закономерности, выявленные на мышах и обезьянах в рамках доклинических исследований вакцины, воспроизводятся на 1 и 2 фазах клинического испытания, так этого нельзя, если наука настоящая. Необходимо, чтобы 1 и 2 фазы (о третьей уж молчим) фальсифицировали-опровергли предыдущие опыты.

Подобные изощрения — широкое поле для сатиры и юмора.

В материалах того же форума S. Greenland предположил, что могут иметься различия в стандартах для данных и правил вывода, соответствующих научному продвижению, с одной стороны, и информированию для принятия решения в области здравоохранения, с другой” [120].

Наконец, в работе 1988 г., S.F. Lanes поставил некую точку (курсив мой. — А.К.): “Следование индукции заставляет ученого взять на себя ответственность за принятие решения, когда доказательств достаточно, чтобы обеспечить основу для действий. Но причинную связь никогда нельзя установить точно, *поэтому не может быть никаких научных оснований для принятия политических решений*. Может иметься только убеждение, что теория причинности верна. Следовательно, индуктивистский принцип, которым руководствуются при применении научных теорий, подразумевает утверждение, что политические решения требуют субъективного одобрения ученых. Индукция установила ошибочную границу между наукой и политикой” [121].

“Публичный имидж ученых был искажен их участием в формировании государственной политики...” [121].

“Хотя мы все хотели бы, чтобы научное знание использовалось с пользой, у ученых иные цели,

чем у практиков. Наука настойчиво ищет истину, но иногда непроверенные и даже опровергнутые теории подходят для действия” [121].

И, более недавно:

“Причинный вывод неявно, а иногда и явно встроен в практику общественного здравоохранения и в формулирование политики” (2013) [22].

Согласно источнику USEPA-2014 [159], “ведущий современный эпидемиолог” (см. выше цитаты) К.К. Rothman одобрял в 2005 г. предложение только представлять научные данные, оставив определение причинности для политиков [110]. (Мы лишь воспроизводим источник USEPA-2014 [159], поскольку в самой работе [110] такого прямо нет; скорее можно говорить о работе Rothman К.К., 1988 [160], и то между строк.)

Все же резюме здесь одно: есть “чистая наука эпидемиология”, которой занимаются посвященные ученые, и есть более низменная практика, с принятием решений и ответственностью, которой занимаются другие (вспомним приведенную выше притчу о философах и их домочадцах). В первой части настоящего Сообщения 4 нами рассматривалось предложение от 2019 г. добавить к критериям Хилла пункт “Последствие принятия решения” [67]. Авторы этого предложения, вероятно, не относят себя к сторонникам чистой науки. Равно как и М. Susser, придававший много значения именно социальной эпидемиологии [10].

В сообщении НКДАР-2012 (опубликовано в 2015 г.) [161] есть специальный раздел, в котором рассматриваются возможные шаги в том случае, когда гипотезы не могут быть проверены или же не были проверены либо были опровергнуты. Возможности следующие (курсив мой. — А.К.).

1. “Чисто научный подход”, как сказано в [161], заключается в исключении подобных гипотез из анализа, но он не может использоваться, когда требуются решения в области здравоохранения. Научный метод, согласно [161] со ссылкой на [162], осуществляется научным сообществом с его особыми *внутренними нормами: правдивостью, последовательностью, согласованностью, проверяемостью, воспроизводимостью, валидностью, надежность, открытость, беспристрастность и прозрачность*³⁷.

2. Создание условных прогнозов (т.е. основанных на гипотезах, которые в настоящее время не считаются доказанными) для оценки риска. Интерпретируя доступную научную информацию, исследователи могут осуществлять экстраполяции и оценки, используя научно обоснованные модели и правдоподобные предположения, и характеризовать степень неопределенности (про последнюю в радиационных дисциплинах см. в [65]). Такие научно обоснованные оценки могут

использоваться *для информирования* в процессе принятия решений³⁸.

3. Подход на основе ненаучных проблем, который может учитывать, а может и не учитывать научные выводы. В подобном случае лица, принимающие решения, основывают соображения на *нормах, внешних по отношению к науке: социальной ответственности, этике, полезности, осмотрительности, осторожности и практичности применения*. Данные соображения, хотя и важны, выйдут за рамки компетенции НКДАР и упоминались в [161], как там сказано, “только для полноты картины”³⁹.

В целом трудно представить себе “чистую науку радиационную эпидемиологию” (в отличие от радиобиологии), хотя дискуссии о типе зависимости “доза — эффект” для канцерогенных последствий облучения в области малых доз (до 100 мГр излучения с низкой ЛПЭ) — линейной беспороговой, сублинейной, линейно-квадратичной, линейной с порогом или гормезисной (что реально недоказуемо — НКДАР, BEIR, NCRP [118, 157, 161, 164]), может, и следует отнести к “чистой науке”.

Сложность медико-биологических связей не поддается определению какими-либо правилами/критериями

Согласно Hofler M., 2005 [51], эвристическая ценность критериев Хилла стремится к нулю по мере увеличения сложности причинной системы и повышения уровня неопределенности для причинности. Сложные причинно-следственные системы включают множество гипотез и предположений о смещениях (bias). Если сложность (комплексность) становится очень большой, то мнение о том, следует ли применять соображения Хилла, приближается к решению путем подбрасывания монеты.

Названное положение повторено в Phillips C.V., Goodman K.J., 2006 [165]: сложные системы не поддаются простым практическим правилам.

Следует отметить, что в качестве основного подхода к оценке причинности в [51, 165] предлагается контрафактический [7, 9], т.е. критерий “Эксперимент контрафактический”, который А.В. Hill когда-то рассматривал как “самую сильную поддержку гипотезы причинности” [11]. Но даже такой сильный критерий (“предотвратимость” и пр. [7]) может не работать в сложных системах: устранение одной причины не обязательно устраняет другие причины, которые могут приводить к тому же исходу. Только при дополнительном ограничении, что никакая другая *достаточная* причина не действует, будет обеспечена дедуктивная связь между причинной гипотезой и предотвратимостью [76].

Согласно современной эпидемиологии, множество актуальных воздействий имеют комплексную причинность из необходимых и достаточных компонентов К. J. Rothman [1, 13, 53, 110, 137, 138, 143, 160], потому опасения из [76] актуальны. Что же касается радиационного фактора, то он не может выступать как необходимый компонент “созвездия” достаточной причины лучевых эффектов, поскольку для таковых, как правило, отсутствует специфичность. Если для детерминированных эффектов радиации и можно придумать некую специфичность (лучевой ожог, лучевой синдром), то для стохастических последствий облучения это невозможно [1, 8].

Мозаичность критериев вследствие их мультидисциплинарности. Недопустима никакая интеграция разнородных данных для оценок причинности

Этот крайний подход экстравагантного [6] теоретика медицины В. G. Charlton (Великобритания) из его часто цитируемой работы 1996 г. [48] подробно рассматривался нами ранее [6].

Согласно указанному автору, мультидисциплинарный и междисциплинарный подходы в эпидемиологии приводят к мозаичности. Каждый элемент мозаики, состоящий из конкретных данных, может быть оценен на научную валидность, но метод, которым эти элементы связаны-склеены, не является научным, потому что, согласно [166, 167], комбинация доказательств от разных дисциплин не может быть самостоятельной научной дисциплиной. Если пробелы в доказательствах для одной дисциплины заменяются или обходятся с помощью данных из другой дисциплины и наоборот [168], то получается, что совмещается малосовместимое. Действительно, поскольку интегрируются данные, полученные путем ряда “несоизмеримых” (или качественно различных) подходов, методологий и способов доказательности, то эпидемиология должна положиться на суждение, на “здравый смысл”, в большей степени, чем другие науки [48].

В. G. Charlton сравнивает доказательные свидетельства в эпидемиологии с “сетью” связанных данных из множества дисциплин. Недостаточность или некорректность в какой-то ячейке не “рвет”, не устраняет всю сеть, а только ослабляет ее, ибо это *сеть*. То есть эпидемиологические гипотезы весьма слабо фальсифицируемы (согласно положениям К. Поппера [126]): противоречивые результаты могут только изменить равновесную вероятность мультифакториальной причинной обусловленности. Это объясняет длительную жизнь и устойчивость иных эпидемиологических гипотез даже перед лицом противоречащих фактов. Напротив, “настоящие” научные гипотезы построены на *цепочке* свидетельств и отдельных доказательных звеньях, а потому устранение лю-

бого звена как некорректного сразу же полностью устраняет и саму гипотезу [48].

О том, что интеграция информации может снижать ценность и важность критериев, говорили и другие авторы, однако исходя уже из иного, более субъективного аргумента [66]. В рамках подобной интеграции исследователи могут интерпретировать критерий в зависимости от того, каким образом он *соответствует доступным данным*, а не определять, *соответствуют ли данные критерию*. Отсюда делается вывод, что такие моменты должны быть проанализированы для каждой уникальной ситуации [66].

В целом приведенные рассуждения не лишены смысла, когда речь идет не об эпидемиологии, а например, о психологии или, скажем, о непроверяемых на практике ретроспективных дисциплинах. Либо о ниспровергающих или, наоборот, “открывающих” теориях типа “контактов” Э. Денникена. Но в случае подхода к оценке причинности преимущественно в качественном смысле, предусматривающего “вес свидетельств” (WoE), столь абсолютистские рассуждения делают бессмысленными многие практические шаги международных организаций, направленные на сохранение здоровья и окружающей среды: МАИР [114, 115], USEPA [45, 61, 104–109] и др. [28]. Кто-то скажет, вновь, что это “прагматизм”, а не наука, но при ином, “строго научном” подходе, скорее всего, не останется никаких реальных обсервационных дисциплин.

Экспертный субъективизм

Хотя целый ряд методологий по оценке WoE для установления причинных связей предусматривает, как отдельно, так и в комплексе с иными подходами, экспертные заключения [28], у некоторых авторов акцент на суждениях для критериев причинности “вызывает беспокойство”, поскольку “субъективные методы в науке” должны применяться как можно реже (Burch P.R., 1983 [169]). В определенном смысле это верное заключение, если не придавать ему оттенок абсолюта. Скажем, даже при применении одного и того же комплекса критериев к одному и тому же эффекту могут иметься разночтения в выводах у разных исследователей. Выше упоминался подобный пример связи между употреблением алкоголя и раком молочной железы, разобранный в [43, 91]⁴⁰.

“Классический подход уязвим для субъективности при оценке доказательств и манипулировании доказательствами, и заинтересованные стороны, потенциально затронутые выводом о том, что связь является или не является причинной, могут занять противоположные позиции по интерпретации доказательств” [22].

Как сказано в [173], использование комплекса критериев, порой, реализуется в цитирование ре-

зультатов “за” или “против” причинной гипотезы, сводясь только к профессиональному суждению. В исследовании [43] и нами ранее [2, 6, 8, 10] указывалось, что применение разными авторами девяти критериев Хилла не отличается унификацией: то или иное положение может опускаться без всяких объяснений, в результате число критериев, а также порядок их значимости, сильно варьируют в зависимости от понятий и суждений исследователей.

“Если метод установления общей причинности позволяет аналитику решать *ad hoc*, какие критерии необходимы для данного набора данных *posteriori*, результатом является не что иное, как личное мнение, маскирующееся под философски обоснованное, объективное толкование” [34].

“Точки зрения (viewpoints) для суждения”, как рассматривал А.В. Hill критерии, представляют собой “тонкий переход к субъективизму”. “Принимая субъективную интерпретацию вероятности, мы переносим объект интерпретации с наших теорий о причинности патологии на наши личные убеждения” [123].

“Причинные критерии и логика полагаются на наилучшее профессиональное суждение (Best Professional Judgment; BPJ), чтобы синтезировать линии доказательства” [105].

“Качественные критерии подвержены когнитивным ошибкам” [20].

Есть критика подобного рода и у К.В. Rothman [53]: “Контрольные списки (checklists) придают обманчивый вид и бессмысленный авторитет изначально несовершенному и творческому процессу”.

Согласиться со столь узкой трактовкой вряд ли возможно. Еще в 1983 г. А.М. Lilienfeld в ответ на подобные претензии указал, что вопрос суждения, который выходит за рамки статистической вероятности, относится не к критериям, а к науке вообще. “Можно применить вероятность к результатам конкретного эксперимента или обсервационного исследования, но невозможно установить вероятностное утверждение и доверительные ограничения на истину научной гипотезы. Основным здесь является то, что демонстрация причинности зависит от синтеза наборов данных разных типов” [129]⁴¹.

Даже величина доверительных интервалов — это дедуктивный выбор [174], т.е. вопрос суждения.

Дальнейшая критика критериев причинности более узка и касается уже не концептуальных проблем, а конкретных моментов, нередко субъективных.

Критерии причинности нечетко описаны, порой совпадают и даже противоречат друг другу

“Нечетко описаны и совпадают” — согласно источнику [82] со ссылкой на работы [175–177], среди которых две первые нам недоступны, а в

пособии [177] ничего подобного не обнаружено, равно как и в его третьем издании [59]. Относительно претензии “совпадают” — уже рассматривалось в [3, 6–9].

“Конфликты критериев” нашел М. Susser [41], и это повторили за ним в обзоре [49]. Сказано [41], что сугубо специфичная причинная связь, “которая приближается один к одному”, нередко является искусственным творением. Артефакт следует из обнаружения эффектов какого-либо агента и последующей реклассификации эффектов по отношению к указанному агенту, т.е. задним числом. Выходит, что конкретность является следствием, а не предшествующим фактом обнаружения причины: причина привела к критерию, или сам критерий создал причину? Иными словами, М. Susser, озвучив “обратную причинность” [5, 8], полагает, что здесь критерий “Специфичность” противоречит критерию “Временная зависимость”.

Аргумент и пример довольно замысловатые и приведены нами как бы “для коллекции”. Хотя формально придумать случаи конфликтов критериев можно немеренно, если взять, к примеру, пункты “Биологическое правдоподобие” и “Согласованность” с текущими фактами и теоретическими знаниями”, с одной стороны, и эпидемиологические ассоциации — с другой. О том, что имеются реальные, но биологически неправдоподобные зависимости, равно как и наоборот, а также о том, что нередко RCT противоречат результатам эпидемиологических исследований, отмечалось нами ранее [6, 9].

В обзоре [49] все это изложено, вновь [6, 10], в оруэлловском стиле: “Возможность того, что причинность подтверждается нарушением причинного критерия, или выполнение критерия для уменьшения причинности”. Речь идет, в частности, о разобранном нами в [8] негативном контроле при тестировании причинности, основанном на отсутствии специфичности⁴².

Критерии — только для ассоциаций в эпидемиологии

“Ученые в наши дни склонны придерживаться вежливой фикции, что все науки равны. ...мы говорим так, как будто область и методы исследования каждого ученого [в его области] так же хороши, как и у всех других ученых, и, возможно, немного лучше. Это заставляет всех нас сердечно рекомендовать друг другу правительственные субсидии”.

Platt J.R., 1964 [179]⁴³

На основе изложенного нами ранее [1–10], это представляется как бы самоочевидным, но, по-

рой специально подчеркивается некоторыми авторами [33, 51, 173, 180]. Однако критерии Хилла используются для определения как индивидуальной причинности (что уже не эпидемиология, а медицина) [10], так и для экспериментов на животных [9, 10].

Несмотря на практическое применение, есть сомнения в том, насколько адекватно критерии Хилла способны тестировать гипотезы в экспериментальных исследованиях, например по оценке МОА — “Способа действия” фактора — в экологии, экоэпидемиологии и экотоксикологии (см. прим. 22) [180].

Но — тестируют ведь [9, 10].

Нет проверки качества исходных данных, на основе которых применяются критерии

В 1979 г. записной критик и диспутант, доньяне большой авторитет в эпидемиологии и ЕВМ, A.R. Feinstein⁴⁴, предъявил свои претензии к критериям Хилла [79, 80] (частично мы рассматривали это ранее [6, 9]). Главным является то, что исследования различного эпидемиологического дизайна поставляют “входные данные в черные ящики” критериев, и нет отдельных критериев качества ни данных, ни используемых методов. Якобы в отсутствие четко сформулированных научных стандартов (при введении в “черный ящик”), использование критериев предполагает, что решения о причинной связи требуют только набора чисел, которые имеют правильный размер и правдоподобные коннотации (connotations) [79, 80].

Со ссылкой на пособие Rothman K.J., Greenland S., 1998 [138], в работе Gori G.V., 2004 [140] развиваются сходные положения. Сказано, что удивительно, почему эпидемиологи не включили в число причинных критериев основания для уверенности в эпидемиологических исследованиях, а именно: что измеренное — это то, что действительно собирались измерить, причем с необходимой точностью и статистической ошибкой, и что систематические уклоны (bias) и конфаундеры контролировались в максимально возможной и достаточной степени.

В 2015 г. аналогичные мысли про необходимость контроля за качеством данных были высказаны и другими авторами, при использовании адаптированных критериев Хилла для экоэпидемиологии и экотоксикологии [186].

В первой части настоящего Сообщения 4 [10] и здесь выше было видно, что некоторые авторы попытались ответить на подобные претензии:

- Ввели критерий “Отсутствие альтернативных объяснений” (т.е. уклонов и конфаундеров) [10];
- Добавили к критериям иерархию доказательности дизайнов в областях ЕВМ и эпидемиологии [102] (см. выше и табл. 3);

- Расширили список критериев включением систематического обзора и мета-анализа [90].

Насколько все это целесообразно, предоставим судить читателю. Поскольку, начиная от Feinstein A.R., 1979 [79, 80], озвучивались какие-то сугубые банальности. Каждому ясно, что исходные данные должны быть корректными и корректно полученными.

Нет экспериментальной проверки действенности критериев

“В контролируемом испытании, как и во всех экспериментальных работах, нет необходимости в [таком] поиске точности, чтобы выбрасывать из окна здравый смысл”.

Hill A.B., 1963 [116]⁴⁵

Как указано в [22], в идеальном мире установление причинности следовало бы выполнять с помощью RCT, на которых и должны были бы основываться решения в области здравоохранения и социальной политики. Но ясно, что RCT часто неэтичны, непрактичны или слишком продолжительны [22, 118, 183–185] для базирования на них своевременного принятия решений. Поэтому используются обсервационные, эпидемиологические исследования и критерии причинности, чему и посвящен весь цикл наших сообщений [1–10].

Но насколько использование критериев Хилла для установления причинности ассоциации работает на деле? Проверяли ли этот момент и как?

Ясно, что кроме “Временной зависимости” наиболее адекватно будет работать критерий “Эксперимент контрафактический”, т.е. наблюдение за эффектом, когда его гипотетическая причина по тем или иным причинам (естественным или так называемым “манипулятивным” [26]) ослаблена либо устранена [5, 8]. Есть даже предложение свести все критерии причинности преимущественно к манипулятивным подходам [26].

Вопросом проверки действенности критериев Хилла озаботились еще в 1988 г. [123, 148], хотя и отмечалось, что не имеется методических стандартов, с которыми их применение можно было бы сравнивать в эпидемиологии [123]. В 2004–2005 гг. ранее цитируемые нами [2, 6, 8–10] C.V. Phillips и K.J. Goodman, отрицающие применение руководящих правил каузальности, заявили, что, по их данным, валидационной проверки причинных критериев не проводилось [18, 33]. Позже, в 2007 г., сходное мнение было высказано и про применение причинных критериев для оценки экологического риска [20].

Вопросы, таким образом, ставились уже давно (1988–2007), и они так и остались в литературе. Но выше было видно, что в работе Swaen G., van

Amelvoort L., 2009 [55], во-первых, была разработана методика рангового взвешивания критериев Хилла, и, во-вторых, проводилась ее проверка путем оценки вероятности канцерогенных эффектов для 159 агентов, выбранных по списку МАИР (IARC) в категориях 1 и 2А (т.е. с наибольшими рисками). Частота согласованных (точных) прогнозов составила 82%. Модель верно классифицировала 89% канцерогенов категории 1 и 71% канцерогенов категории 2А.

И можно было бы сказать, что наконец-то проведена экспериментальная проверка применения критериев Хилла, если бы не возможности “круговой логики” [112, 113]. Ведь МАИР часто использует для подтверждения выводов об опасности соединений именно критерии Хилла [114, 115].

В то же время можно вспомнить значительный пласт работ по сравнению одних и тех же медико-биологических эффектов, выявленных в наблюдательных исследованиях и в РСТ. Хотя в отдельных работах продемонстрирована гетерогенность результатов сравнения (эффекты больше – меньше – равновелики; сводки см., например, в [187–189]), в Кокрейновском систематическом обзоре от 2000 г. [187] интегрирующим заключением было то, что средние результаты наблюдательных исследований оказались “удивительно сходны” с результатами РСТ. Есть и другие исследования с аналогичными выводами [188], в том числе мета-анализ мета-анализов [189]⁴⁶.

Критерии “обременены оговорками и исключениями”

“А если бы он вез патроны?”

Кино “Невыдуманная история” (1963)

Представленная в эпиграфе сентенция уже использовалась нами ранее [6, 7, 10], но не было обнаружено ничего лучше для отображения сути ситуации, рассмотренной в настоящем подразделе. Поэтому мы повторяем данную фразу опять. Согласно Hill A.V., 1966 [190], где процитирован A. Gide, “все было сказано раньше, но, поскольку никто не слушал, сказанное следует повторить”.

“Эти критерии, предложенные Хиллом, обременены оговорками и исключениями”. Данная фраза, на которой основывается отрицание каких-либо правил причинности, согласно [121], имела место уже в работе Rothman K.J., 1982 [191] и дожила почти в таком виде до публикаций 2005 г. [52, 110] и до 2–3 издания “Modern Epidemiology” от 1998 г. и 2008 г. [13, 138]⁴⁷. Следует отметить, что в последнем пособии по эпидемиологии Rothman K.J., 2012 [53] резюмирующие акценты смягчены и, хотя пригодность критериев Хилла для причинного вывода отрицается, признается целесообразность их применения при проверке гипотез⁴⁸.

Ранее нами указывалось [6], что K.J. Rothman и S. Greenland для каждого критерия Хилла, включая необходимость самой ассоциации [1, 3, 5, 6, 10], находили некое исключение, порой головоломное. В результате получалось так, что ни один критерий (кроме “Временной зависимости”) не является даже слабым “пунктом”, позволяющим иметь какие-то суждения [13, 52, 53, 110, 137, 138].

Фразы или конкретный материал об исключениях, не подтверждающих тот или иной критерий, рассыпаны в работах разных авторов. Даже для пяти критериев из Сообщений Главного врача США о последствиях курения от 1964 г. и 1982 г. [31, 32, 192] почти каждым диссидентом, отрицавшим связь между курением и раком легкого (см. в [2, 10]), предлагались свои “патроны”, опровергающие какой-либо руководящий принцип [2, 169, 193]. То же – для девяти критериев Хилла [18, 26, 51, 72, 78, 82, 86, 121, 139, 165, 194, 195].

Попытки изо всех сил найти исключения к отдельным пунктам, чтобы опровергнуть все в целом, иной раз напоминают начетничество. Это, как если бы кто отрицал пятиборье потому, что для бега всегда может найтись в виде исключения безногий, для стрельбы – плохо видящий, для фехтования – безрукий, а для плавания и конного – не умеющий плавать и скакать. Или тоже безногий. Особо отличился здесь такой авторитет эпидемиологии, как K.J. Rothman, если не считать диссидентов от вреда курения.

Ранее нами разбирались подобные исключения по каждому критерию в [1, 3, 6–10], хотя и не все. Некоторые примеры похожи на пародии и, потому, заслуживают приведения.

• “Ассоциация”. Стандартизованная частота смертности от рака легкого у финских мужчин в 1960–1961 гг. была вдвое выше, чем у белых мужчин в США, тогда как потребление сигарет в 1950 г. в Финляндии было вдвое меньше, чем в США. Еще вопрос: если канцерогены табачного дыма – универсальны, то почему одни типы рака положительно связаны с курением, а другие, хотя таких и мало, – отрицательно? И многое другое в этом роде (Burch P.R., 1983 [169]).

• “Сила связи”. “Большие эффекты почти всегда связаны с ошибками и смещениями (bias)”. Эти рассуждения из Ioannidis J.P., 2016 [195], повторенные в Cox L., 2018 [26], альтернативны эпидемиологическим основам причинности [3, 8, 10], да и логике.

• “Постоянство ассоциации”. Невыполнение критерия не исключает причинной связи: переливание крови может вызвать ВИЧ-инфекцию, но оно не постоянно это делает, поскольку вирус также должен присутствовать (Rothman K.J., Greenland S., 2005 [52, 110]; Rothman K.J. et al., 2008 [13]). Чтобы придумать подобный аргумент

против воспроизведения данных, надо, конечно, иметь неординарное мышление.

- “Временная зависимость”. Причина *C* предшествует предполагаемому эффекту *D*. Однако обратный порядок времени не является доказательством против гипотезы, что *C* может привести к *D* (Rothman K.J., Greenland S., 2005 [52, 110]; Rothman K.J. et al., 2008 [13]). Без комментариев.

- “Биологический градиент”. “Витамин Е демонстрирует эффект двуликого Януса”: порядка 70% исследованных доз снижают смертность, а порядка 30% — увеличивают ее (Ioannidis J.P., 2016 [195]). Имеется в виду — какая тогда получается зависимость “доза — эффект”? Еще [169]. Курильщики можно разделить на две категории: социальные и привычные, конституционные. Социальные курильщики, как правило, являются легкими по пристрастию и с готовностью бросают курить; привычные, *генетически предрасположенные*, как правило, являются заядлыми курильщиками. Следовательно, в группе легких курильщиков, с малым числом сигарет, будут преобладать социальные курильщики, и связь с раком легкого будет относительно слабой, а в группе заядлых курильщиков, курящих много, будут преобладать привычные, генетически предрасположенные, и связь с раком легкого будет сильной. Отсюда, в целом по популяции, и наблюдается соотношение “доза — эффект” (Burch P.R., 1983 [169]). То есть связь между числом сигарет и смертностью от рака легкого, достигающая $RR = 40-60$ [3, 31, 73, 192], получается “кажущейся”.

- “Биологическое правдоподобие” и “Согласованность”. Много примеров нами приводилось в [6, 9]. Вот еще, порельефнее. Предлагается ответить на следующие вопросы: а) “Если наблюдаемая ассоциация соответствует существующим знаниям, присвоили бы вы ей меньший вес (по сравнению с весом существующих знаний), если бы ваши наблюдения не соответствовали существующим знаниям?” и б) “Если наблюдаемая ассоциация не соответствует существующим знаниям, присвоили бы вы ей более высокий вес (по сравнению с весом существующих знаний), если бы ваши наблюдения соответствовали существующим знаниям?” (Cox D.R., Wermuth N., 1996 [196]; цитировано по [51]).

- “Эксперимент контрафактический”. Все исследования по снижению частоты смертности от рака легкого при отказе от курения касаются бросивших курить самостоятельно. Конституционная гипотеза предсказывает, что бросившие курить будут, как правило, не привычными, заядлыми курильщиками. Исследования оставшихся курение до того, как они это сделали, по сравнению с продолжающими курить, показывают, что в среднем это разные группы населения со многими различными чертами (Burch P.R., 1983 [169]).

- “Аналогия”. Связана с изобретательским воображением исследователей, которые могут найти аналогии всюду (Rothman K.J., Greenland S., 2005 [52, 110]; Rothman K.J. et al., 2008 [13]). Между тем метод молекулярной аналогии — один из подходов при оценке канцерогенности или токсичности неизвестного соединения МАИР, токсикологическими и экологическими организациями [9].

Вероятно, можно добавить и еще примеры, хотя мы, похоже, здесь и ранее [1, 3, 6–10], извлекли все близко к возможному. Изобретательность впечатляет, но возникает мысль: если восемь из девяти критериев (хотя и для “Временной зависимости”, как видно выше, не все однозначно) неабсолютны по отдельности, пробовал ли кто-то найти такую причинную ассоциацию, которая не укладывалась бы во все восемь пунктов? Этого вопроса мы не нашли ни у одного автора сотен публикаций за около 60 лет.

Тем не менее нечто близкое изложено в Guzelian P.S., 2005 [34], где критикуется приведенная выше критика: “Утверждения о том, что каждый из критериев Хилла можно обойти (обычно изречаемые авторитетно, без подтверждений), оказываются нерасчетливыми (improvident), если целью является установление причинности”⁴⁹.

Наконец, рассмотрим, порой кратко, что же предлагается взамен критериев тогда, когда поставить прямой эксперимент невозможно.

МЕТОДЫ ОЦЕНКИ КАУЗАЛЬНОСТИ ВНЕ ПРИЧИННЫХ КРИТЕРИЕВ

Помимо критериев причинности, в эпидемиологии (разных направлений) известны иные каузальные модели и подходы, ставшие относительно известными только в последние десятилетия. Насколько они востребованы в реальных исследованиях и при формировании практических рекомендаций — вопрос иной; мы его уже отчасти рассматривали [1] и рассмотрим еще. Можно, как сказано выше, следовать “последней эпидемиологической моде”, отринув “критерии великого светила прошлого [A.V. Hill], напоминающие “научные” методы Темных веков (Dark Ages)” [33], но насколько это реально? Ведь от качественного, иллюстративного подхода до реально дела путь не близкий. На это указывал еще M. Susser, рассматривая “Popperian Epidemiology” [197] (см. в [9]; прим. 36).

Есть источники, в которых был выполнен обзор, использованы или просто перечислены известные модели причинности в эпидемиологии [12, 13, 19, 24, 35, 67, 133, 139, 198–207]. Помимо англоязычных, имеется польский (2009) [204], французский (2014) [205] и японские (2009–2018) [206, 207] (и др.) обзоры. Для русскоязычных ис-

Компонентные достаточные причины эффекта

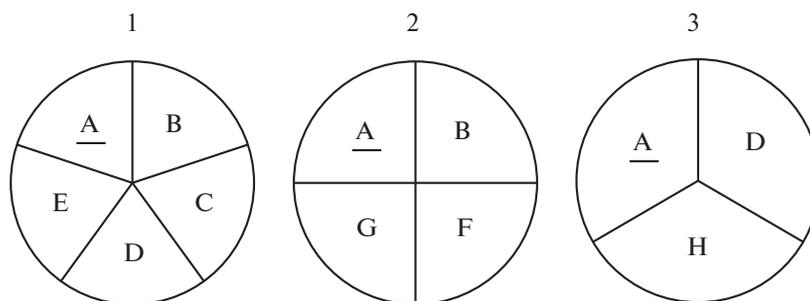


Рис. 1. Модель SCC (Sufficient-Component Cause model), или “Causal Pie”, т.е. “Пирог причинности” К.Ж. Rothman (составляющие достаточной компонентной причины эффекта). Наша модификация [1] исходной схемы (иное число секторов сравнительно с оригиналами [13, 53, 110, 137, 138, 143, 216, 217]).

Fig. 1. Model SCC (Sufficient-Component Cause model), or “Causal Pie”, that is, K.J. Rothman (constituents a sufficient component cause of the effect). Our modification [1] of the original scheme (a different number of sectors compared to the originals [13, 53, 110, 137, 138, 143, 216, 217]).

точников ничего подобного обнаружено не было, причем не только для публикаций по эпидемиологии, включая все известные нам пособия по предмету. Поэтому, как бы ни были, порой, кратки и/или недостаточны в формализованном плане приведенные ниже данные (кроме сведений про модель SCC), представленный нами материал, вероятно, является первым обобщающим обзором на русском языке. Насколько это нужно при сложившемся положении вещей, сказать трудно.

В целом каузальные подходы в эпидемиологии вписываются в следующий перечень причинности для этиологии заболеваний [208]: вероятностная, детерминированная и смешанная. Критерии причинности, что не раз нами рассматривалось [1–10], входят в первую категорию.

Модель достаточной компонентной причины К.Ж. Rothman: детерминизм

Исторические вехи

Оригинальное наименование: “Sufficient-Component Cause model of Rothman” или “Модель SCC”. Эта модель отчасти разбиралась нами в обзоре [1], при обсуждении типов причин в медико-биологических дисциплинах. Некоторые аспекты, в том числе декларативная значимость положений К.Ж. Rothman для современной эпидемиологии, были затронуты здесь выше.

Корни понятия о многопричинности явлений явно уходят в глубь веков, но в качестве исходного источника всюду [13, 133, 139, 143, 208] называются положения J.S. Mill (1806–1873; Англия) от 1843 г., согласно которым реальная причина эффекта (достаточное условие) является совокупностью ряда условий (необходимых, но недостаточных по отдельности) [209]. Далее эти понятия

разрабатывались философом J.L. Mackie (1917–1981; Великобритания) с его известной концепцией INUS (“Insufficient and Non-redundant part of an Unnecessary but Sufficient condition”) в статье 1965 г. [210] и, затем, в монографии 1974 г. [211]⁵⁰. (Кроме того, идея о комплексности и взаимодействии различных причин для единого следствия была выдвинута в работе Lewis D., 1973 [214], также в философском аспекте; эта временная веха упомянута в [51]).

Подход J.L. Mackie к причинам, связанным с состояниями INUS, не был разработан специально для эпидемиологических целей, но оказалось, что он имеет много общего с методами данной дисциплины [21].

Суть модели SCC К.Ж. Rothman и примеры

В 1976 г. К.Ж. Rothman, будучи в возрасте 31 года, впервые представил свою графическую модель SCC [215], которая позже воспроизводилась в его (как правило, вместе с S. Greenland и, порой, другими соавторами) публикациях и монографиях, вплоть до минимум 2014 г. (когда К.Ж. Rothman было 69 лет) [13, 53, 110, 137, 138, 143, 216, 217] (рис. 1). Та или иная модификация диаграммы (“Causal Pie” – “пирог причинности”, более похожий на торт) приводится почти во всех современных пособиях и обзорах по эпидемиологии [21, 24, 27, 40, 56, 57, 59, 177] (и др.; перечислены только ранее использованные источники).

Концепция основывается на том, что большинство причин являются компонентами, составляющими достаточные причины, но сами по себе они недостаточны. К.Ж. Rothman упоминает “созвездие (constellation) явлений (на рис. 1 – сектора), которое является достаточной причиной (на рис. 1 – круги). Причем и все “созвез-

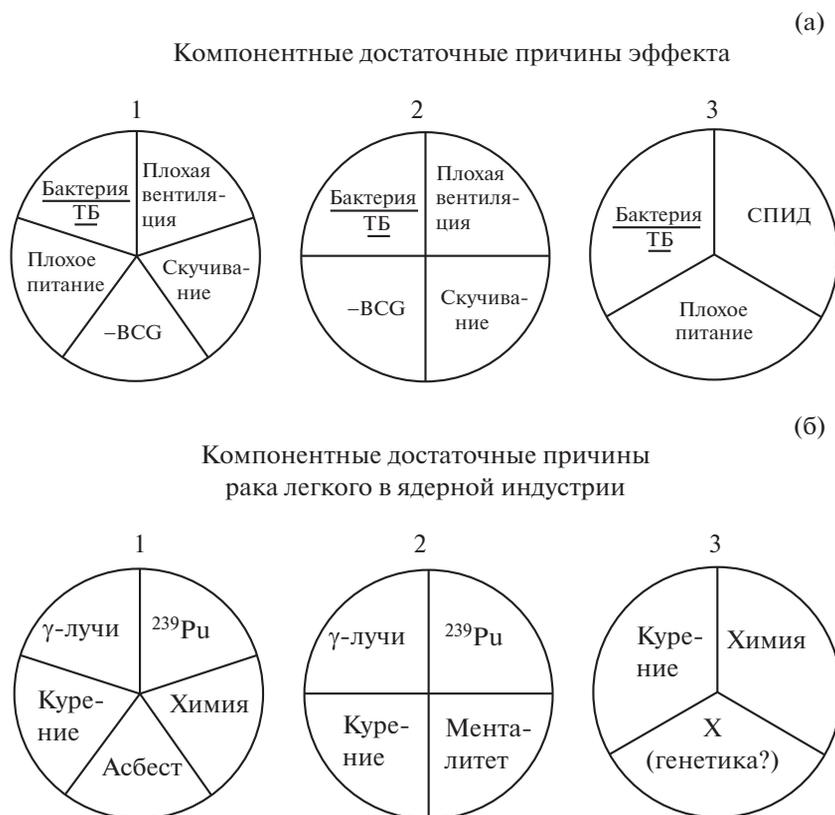


Рис. 2. Модели SCC для трех компонентных достаточных причин: а – туберкулеза [1]; б – рака легкого у работников ядерной индустрии (подробности в тексте).

Fig. 2. SCC models for three component sufficient causes: a – tuberculosis [1]; b – lung cancer in nuclear workers (see text for details).

дие”, и его компоненты, называются “причинами”. В результате термин “причина” сам по себе не указывает, будет ли явление достаточной причиной, или же компонентом достаточной причины [110, 215].

Из рис. 1 видно, что один компонент – “А” – является *необходимым*: он входит во все причинные созвездия. На рис. 2, а представлен такой пример [1]: инфицирование бактерией туберкулеза входит во все достаточные компонентные причины данного заболевания, вкуче с иными, уже не необходимыми, но достаточными, компонентами.

Ранее [1] нами рассматривался вопрос: “Может ли радиация быть необходимым компонентом “созвездия” достаточной причины тех или иных лучевых эффектов?” Судя по всему, нет, ибо известно, что для радиационных эффектов на подобных уровнях отсутствует специфичность. Нет такой патологии или изменения, которое индуцировалось бы исключительно облучением, но не каким-то иным фактором (хотя бы активными формами кислорода), равно как нет и соответствующего лучевого биомаркера [65, 118, 157, 161]. Некоторым исключением могут являться лучевые

ожоги, которые имеют отличия в своем развитии, терапии и прогнозе от других ожогов [161, 218]; иным исключением будет, вероятно, лучевой синдром [218].

В обзоре [1] отсутствие лучевой специфичности эффектов иллюстрировалось нами схемами компонентных достаточных причин радиогенных раков (сходную модель можно построить для лейкозов) и лучевых сердечно-сосудистых патологий. В настоящей публикации мы представляем модель SCC для рака легкого у работников ядерной индустрии (рис. 2, б).

Ядерные работники могут подвергаться воздействию внешнего облучения, инкорпорированных радионуклидов (в частности, ^{239}Pu), асбеста, химических соединений, а также курения [219, 220]. Все эти агенты способны вызывать рак легкого, в том числе в комбинациях друг с другом [118, 221]. Но ни один из них, включая внешнее и внутреннее облучение, не является *необходимой* причиной названной патологии, что иллюстрирует “созвездия” на рис. 2, б. Более того, некоторые “достаточные” причины, сами по себе, возможно, индифферентные, способны влиять на эффект других компонентов, как это видно для

второго круга-“созвездия”. Может наблюдаться даже синергия [215] (пример: воздействие радона и курения [65, 118, 161, 221]). Или модуляция: менталитет, жизненные привычки и установки, могут отражаться не только на интенсивности курения и потребления алкоголя, но даже на накопленной дозе радиации. Так, для японских работников атомной индустрии показана связь между уровнем поглощенной дозы, с одной стороны, и курением, а также алкоголизмом, с другой. Оказалось, к тому же, что для более высоких дозовых групп частота прохождения плановой рентгенодиагностики была ниже [222]. Есть также ассоциация между дозой облучения и частотой смертности от нерадиационно-обусловленных патологий (цирроза печени, рака ротовой полости и глотки, психозов и внешних причин) у работников ядерной индустрии Франции (имитация зависимости “доза-эффект” алкоголем) [223].

Иными словами, менталитет может увеличивать накапливаемую дозу радиации, способствуя, вероятно, согласию на более опасные работы и меньшему следованию нормам безопасности.

Наконец, как видно по третьему кругу-“созвездию” на рис. 2, б, где представлены только нелучевые факторы, не все компоненты достаточной причины могут быть известны. Для рака легкого вероятны и неидентифицируемые генетические причины, хотя ныне выявлены и конкретные гены предрасположенности (например, *CYP1A1 T3801* [224]). Но модель К.Ж. Rothman сильна тем, что для уменьшения неблагоприятного эффекта (“профилактических целей” [27]) порой не обязательно знать все причинные компоненты: достаточно устранить только одну [19, 27, 40, 53, 81, 86, 110, 215]. Скажем, курение – на рис. 2, б видно, что этот фактор входит во все причинные “созвездия”. Можно вспомнить также пример середины XIX в., когда J. Snow устранил эпидемию холеры в Лондоне путем заглушки конкретной водоразборной колонки. Микробная теория патологий тогда еще не была известна (см. в [6, 9]).

Устранение достаточной причины синонимично устранению биологического начала патологии, поскольку таковая начинается, когда последний компонент причины “встает на свое место”. Биологическое начало болезни не обязательно является синонимом появления симптомов или клинического диагноза; последние события могут произойти спустя длительное время после биологического начала. Например, переход от ВИЧ-инфекции к СПИДу (без лечения) может занять более десяти лет [40]. То же самое можно сказать и про развитие радиационного канцерогенеза, модулируемого в течение десятков лет множеством причин [65, 118, 161, 221] (рис. 2, б).

Компонентные факторы, условия или события могут действовать и одновременно, и последовательно, как в случае инициаторов и промоторов рака [27]. Они могут проявляться весьма разнородно друг от друга по времени. Одни факторы могут быть отдаленными, а другие – непосредственными причинами. Например, модель SCC для рака молочной железы включает ген *BRCAl*, ранний возраст менархе, поздние роды и заместительную гормональную терапию во время менопаузы, что охватывает почти всю жизнь [40].

Сила причинного фактора в модели SCC зависит от распространенности (prevalence) дополнительных компонентов в одной и той же достаточной причине. Но эта распространенность часто обусловлена привычками (custom), обстоятельствами или случайностями, не являясь научно обоснованной характеристикой [215].

Модель SCC также иллюстрирует понятие “*сильные*” и “*слабые*” компонентные причины. Компонентная причина, которая может завершить весь комплекс достаточной причины (т.е. 1–3 на рис. 2), но – только вкуче с другими компонентами с *низкой распространенностью*, является “*слабой причиной*”. Наличие подобной компонентной причины приводит к изменению вероятности исхода незначительно, от нуля до значения чуть более нуля, что отражает редкость дополнительных компонент. Но если такие дополнительные причины не редки, а почти повсеместны, то это делает компонентную причину, завершающую вместе с ними весь причинный комплекс, “*сильной*”. Слабая причина приносит лишь небольшое увеличение риска заболевания, тогда как сильная причина значительно увеличит риск [215].

*Модель SCC означает детерминизм.
“Особость” методике К.Ж. Rothman*

Факт детерминистских воззрений К.Ж. Rothman, что редкость для эпидемиологов [35, 42, 133, 141, 144, 146, 225–228], освещался здесь выше, равно как и в предыдущих наших обзорах [1, 6]. Для К.Ж. Rothman (в том числе с соавторами) вероятностная модель только отражает незнание конкретных индивидуальных детерминант риска (2008) [13]:

“По нашему незнанию, лучшее, что мы можем сделать при оценке риска, – это классифицировать людей в соответствии с измеренными показателями риска, а затем назначить средний риск, наблюдаемый в группе, людям в этой группе. По мере расширения знаний или спецификации дополнительных индикаторов риска, его оценки, присвоенные людям, будут отклоняться от среднего значения в зависимости от наличия или от-

сутствия других факторов, предсказывающих результат”⁵¹.

Многие авторы отмечают строгий детерминизм модели SCC [86, 133, 142, 198, 227, 228]: “случайность является эвфемизмом нашего невежества” [225, 227] (и в этом роде [142]).

К. J. Rothman и S. Greenland выводят свою методологию эпидемиологических исследований на основе отрицания индуктивных подходов и руководящих принципов причинности, означающих вероятностную платформу. Но как в этом обсервационном хаосе множества действующих факторов, конфаундеров и смещений (bias) оценить истинную причинность ассоциаций? Если не использовать некие разработанные пункты/принципы/критерии?

Указанные авторы:

1) отрицают первостепенную значимость формальных статистических подходов (“вероятностей”, т.е. значений p или “правдоподобия” – likelihood) [13, 110, 138, 229, 230];

2) придают основное значение тестированию выдвигаемых гипотез [13, 110, 230], т.е. “Popperian Epidemiology”. Насколько можно понять, исследователь должен придумать все возможные компоненты достаточной причины, причем в разных вариантах, и потом тестировать свои гипотезы с помощью различных подходов;

3) особая важность отводится количественной, а не качественной вероятности, т.е. доверительным интервалам, а не значениям p [13, 110, 229, 230]. Акценты ставятся на точности стратификации, измерении ошибки определения, погрешностях и пр. [13, 110, 138, 229]: “Краеугольным камнем научных испытаний является опора на измерения” (Rothman K.J., 1986 [229]);

4) вследствие детерминизма, по-видимому, существует необходимость найти все конкретные компонентные причины и определить их эффекты.

Более нам не удалось понять, что же еще подразумевалось К. J. Rothman и S. Greenland в их методологии, кроме отрицания критериев причинности. По виду перечисленное – стандартные эпидемиологические методы и подходы, но с упором на более точное и углубленное, количественное тестирование изобретенных гипотез:

“Требуется гораздо большее, чем использование списка критериев. Вместо этого следует применять тщательную критику с целью получения количественной оценки общей ошибки, которая затрагивает исследование. Этот тип оценки не тот, который может быть легко выполнен кем-то, у кого нет навыков и подготовки исследователя, знакомого с предметом и применяемыми научными методами. Ни одно из них не может быть легко применено ни судьями в суде, ни также учеными, которые либо не обладают необходимыми

знаниями, либо не имеют времени, чтобы вникнуть в работу” (2005) [110]⁵².

То есть что-то вроде “Бороться и искать, найти и не сдаваться”.

Сказанное, на наш взгляд, несколько замысловато, но банально в плане решения задач обсервационных дисциплин. Хотя, как отмечалось выше, ряд авторов и полагают, что современная эпидемиология основывается на воззрениях и модели К. J. Rothman [19, 86, 142] (статьи 2006–2019 гг.; см. прим. 32–34). Если кто-то лучше понимает особенности “метода К. J. Rothman” (порой с соавторами), то всегда сможет добавить сведения и уточнить наш обзор. А пока, в который раз, остается пространство для конъюнктуры: если никакие причинные критерии, включая даже “Силу связи”, не выполняются, то и не надо: достаточно, чтобы методика включала “правильную стратификацию”, “углубленные оценки” погрешностей, доверительные интервалы и тестирование тех гипотез, которые пришли в голову. То есть нечто и так известное и как бы распространенное. Но – с опорой на мировые авторитеты в эпидемиологии.

Преимущества модели SCC

1. Может идентифицировать мультифакторальную причинность патологии и эффектов [158, 228]. Модель SCC полезна для большинства видов рака, когда как наследственные (генетические), так и приобретенные (экологические) факторы риска играют роль в многоступенчатом процессе канцерогенеза [81]. Попытка оценить долю восприимчивости популяции к некоему воздействию имеет мало смысла, если причинную связь основывать на вероятности, так как роль играет множество факторов и даже случайности [227]. Модель SCC, согласно [228], “отвечает концепции здравого смысла”, объясняя, почему, скажем, кто-то из заядлых курильщиков дожил до 95 лет без рака легкого. Это произошло не только из-за удачи, но и из-за отсутствия одной или нескольких компонентных причин. Здравое понимание данной концепции полезно для политики профилактики заболеваний [228]⁵³.

2. Позволяет выполнять профилактические меры (устранять значительный пул эффекта) путем блокирования даже одного причинного компонента в случае его распространенности [40, 81, 158].

3. Объясняет эффект, когда причинные компоненты в цепочке, как сказано выше, далеко разнесены во времени (отдаленные и непосредственные причины) [40]. Модель SCC адекватна для понимания механизмов индукционного периода [228].

4. Делает понятным различную степень силы связи для разных популяций и в разных условиях,

поскольку сила эффекта для одной компоненты может зависеть от распространенности других компонент [19, 215].

5. Способствует изучению редких заболеваний. Редкие патологии должны иметь редкие причины или нуждаться в большем причинном поле, либо – в определенной последовательности событий, которые происходят редко [228].

6. Делает понятным аддитивность, синергию или антагонизм компонентных причин [228].

Недостатки модели SCC

Критика воззрений К.Ж. Rothman встречается у ряда авторов, как эпидемиологов, так и философов науки, но в большинстве своем присутствует нечеткость, многословность, порой просто “поток сознания”, как это отмечалось нами ранее и для рассуждений вокруг критерия “Биологическое правдоподобие” [6]. Иной раз нелегко понять, что же хотели сказать авторы. Одни находят одно, другие – другое; похоже, что за 45 лет (с первоисточника Rothman К.Ж., 1976 г. [215]) никто не озаботился свести возражения воедино. Мы попытались собрать все и провести систематизацию, хотя и нет уверенности в ее удовлетворительности.

1. При наличии нескольких *достаточных* компонентных причин наблюдается трудность не только при диагностике (поскольку любое “созвездие” может вызвать патологию), но и при оценке воздействия каждого отдельного компонента на популяцию [21, 231].

2. В реальной жизни трудно вычленить действующие на популяцию факторы – компоненты причины, равно как оценить силу их эффекта, если последовательно исповедовать детерминизм. J.N. Morris, автор известной статьи по эпидемиологии от 1955 г. [232] и, затем, одноименной монографии от 1957 г. [233]⁵⁴, в 1988 г. [235] “высоко оценил” первое издание “Modern Epidemiology” под авторством К.Ж. Rothman [137], но отметил, что “как руководство по современной эпидемиологии книга имеет серьезные ограничения”. Указывалось, что главными детерминантами здоровья популяций и подгрупп в них являются экономические, социальные и культурные факторы. Эпидемиологические данные выводятся из исследований популяций и групп как единиц, а не из совокупности индивидуумов с их различными атрибутами [детерминизм]. Это, плюс лабиринт (maze) взаимосвязанных переменных, влияющих на уровень жизни, и множество другого “шума”, часто приводит к неприемлемой неопределенности, а у К.Ж. Rothman нет ответа на такие вопросы [235].

3. Отмечались непрактичность [228] и метафизичность (т.е. отрыв от эмпирического подтвер-

ждения) [225] детерминистской модели SCC. Согласно Davey Smith G., 2019 [142], мысль, что с увеличением знаний можно приблизиться к определенности – это воплощение преувеличенной “персонализированной медицины”, для которой есть множество серьезных, порой непреодолимых ограничений [236, 237]. В частности, для того, чтобы фактор риска давал разрешимость на индивидуальном уровне, величина RR должна достигать порядка 50 [158, 199].

Не имеет значения, считать ли незнание онтологическим (существуют ли действительно непредсказуемые процессы, ведущие к патологии?) или же эпистемологическим (невозможно собрать адекватные данные для определения причинности на индивидуальном уровне) [142]⁵⁵. “Базовые эпидемиологические и биологические рассуждения – это серьезные препятствия для приближения... к “святому Граалю” – нулевым или единичным значениям [индивидуального] риска” [142, 236]. О “мифопоэтике персонализированной медицины” [142] говорилось также выше.

4. В эпидемиологии практически не известны достаточные причины [198, 228]. Наблюдаются просто тенденции к достаточности либо тенденции к необходимости в когортных исследованиях или в исследованиях “случай – контроль” соответственно. Случайность вместо детерминизма играет все более важную роль в эпидемиологии [198]; еще R.A. Fisher отмечал, что биологические процессы часто не становятся детерминированными [239].

Для детерминизма идентификация достаточного компонента подразумевает существование бесчисленных модификаторов, скрывающих эффект на индивидуальном уровне (пример – отсутствие рака легкого у 95-летнего курильщика). В случае сильной достаточной причины, отсутствие эффекта должно предусматривать, что действие указанных модификаторов является более сильным, чем самой причины [133].

5. Мультифакториальная компонентная причинность не может объяснить наличие феномена “доза–эффект” [133]. Действительно, почему должна наблюдаться монотонная зависимость для разных причинных “созвездий” при изменении уровня только одной компоненты? Ведь ее эффект малопредсказуемо опосредуется другими компонентами. В [133] указано следующее: поскольку достаточная причина является полностью достаточной для ее действия, модель SCC не способна легко объяснить, как изменение в количестве компонента (например, в дозе антибиотика) приводит к соответствующему изменению эффекта. Отмечается громоздкость объяснения К.Ж. Rothman [215], предполагающего наличие *набора* достаточных причин-“созвездий”, каждая из которых включает в себя *разную* дозу интересую-

шего компонента (антибиотика). Непонятно, каким образом тогда может фиксироваться плавная, а не дискретная зависимость “доза–эффект” [133] (либо надо предположить наличие как бы бесконечного набора достаточных причин).

6. Модель SCC, о чем не раз упоминалось, не способна объяснить *эпидемиологическую значимость* силы связи [227, 228], каковой принцип, как было видно выше из иерархии критериев Хилла и из предыдущих обзоров [3, 4, 8], является весьма важным для оценки истинности ассоциаций.

7. Модель SCC не обладает предсказуемостью: она не указывает на то, что причинная связь следует предопределенной последовательности. Будет ли данный восприимчивый человек подвергаться воздействию, скажем, табачного дыма в уязвимый для него момент времени, зависит от ряда событий, которые могут быть непредсказуемыми по своей природе. Отсюда имеется сложность в разработке профилактических мероприятий, поскольку модель не всегда способна определить достаточную причину [228]. Это, вероятно, относится в основном к хроническим патологиям. Заметим, что выводы [228] о прогностической слабости модели SCC противоречат заключению работ [40, 81, 158], указывающим на профилактическую способность как преимущество модели (см. в предыдущем подразделе пункт 2).

8. Модель SCC делает нелогичным проведение мета-анализа [228]. Если в разных случаях компонентные причины различаются, и это непредсказуемо, то какой смысл в объединении данных? Между тем, мета-анализ считается важной методикой доказательности как в эпидемиологии [13, 27, 39, 44, 56, 57, 59, 60, 95, 98, 154, 177, 187, 189], так и, особенно, в ЕВМ (“платиновый стандарт” “золотого стандарта” – RCT [240]), и подобные синтетические исследования – прерогатива Кокрейновского общества [44, 94, 95]. Если считать, как некоторые авторы, модель К.Ж. Rothman “идеологией современной эпидемиологии” [19, 86], то без расщепления сознания применительно к мета-анализу тут не обойтись.

Попытки “оправдать” и адаптировать модель SCC

В своей первой публикации 1976 г. [215], К.Ж. Rothman, видимо, осознавая аргументы из предыдущего подраздела, сделал оговорку, что его модель причинности SCC – это некая *мета-физика*, так сказать, “чистая наука”:

“Концептуальные рамки для приведенных здесь [компонентных] причин не предназначены ни для обзора, ни для расширения знаний; [материал] является скорее точкой зрения, устраняющей разрыв между метафизическими представле-

ниями о причине и основными эпидемиологическими параметрами. Таким образом, фокус [рассуждений] не является ни метафизикой, ни эпидемиологией, но – промежуток (gulf) между ними... может облегчить понимание причин патологии”⁵⁶.

Итак, К.Ж. Rothman в самом начале, в 1976 г., попытался осуществить смычку между детерминированной теорией индивидуальной причинности, на деле непознаваемой, и реальной вероятностной популяционной эпидемиологией с ее практическим выходом. Но *ни в одной публикации далее* указанного исследователя (в том числе с соавторами) нам не удалось найти не только намеков на “метафизику” модели SCC, но даже сам этот термин, а подобных публикаций немало, вплоть до 2014 г. [13, 53, 110, 137, 138, 143, 216, 217].

Тем не менее последующие апологеты отвечают критикам модели SCC именно приведенными положениями о “метафизичности”. Мол, К.Ж. Rothman и не собирался внедрять свою модель для практики эпидемиологии. В работе De Vreese L., 2009 [158] сказано следующее⁵⁷:

“...эпидемиологи, похоже, путают практические результаты эпидемиологических исследований на популяционном уровне с метафизическими взглядами на реальность причинности патологий на индивидуальном уровне”.

“С одной стороны, я отвечаю на вопрос о полезном понятии причины для эпидемиологической практики, защищая вероятностную оценку с позиции средних эффектов как подходящую. С другой стороны, я отвечаю на вопрос о метафизических предпосылках причинных понятий в эпидемиологии, защищая взгляд К.Ж. Rothman на причинность как то, что обеспечивает пронципальный (insightful) ответ на этот вопрос”.

Далее автор [158], вновь отдав дань прагматизму, акцентирует внимание на том, что прагматические подходы могут быть ошибочными “с позиции общего и философского анализа”, и, потому, требуются “научные подходы наряду с традиционными”.

Повторяя в 2009 г. [158], что “эпидемиологи должны научиться видеть разницу, которую К.Ж. Rothman точно пытался прояснить, а именно, между метафизическими предпосылками о причинной связи, как это имеет место на индивидуальном уровне, и популяционным подходом в эпидемиологической практике”, L. De Vreese не комментирует то, что сам К.Ж. Rothman ничего подобного, как уже отмечалось, не высказывал с 1976 г. [215]. В [158] есть ссылка в том же смысле и на публикацию Rothman К.Ж., Greenland S., 2005 [110], но в последней ничего об “особости” и “метафизичности” модели SCC не видно.

М. Parascandola, D.L. Weed в 2001 г. [133] попытались осуществить ту самую смычку “практики и метафизики”, когда придали модели SCC, “пирог причинности”, вероятностный характер, приняв, что отдельные компоненты вносят вклад в общую *вероятность* эффекта, а не детерминированы для него [133]. Если один из компонентов отсутствует, то вероятность эффекта уменьшается [133, 158].

*Малая распространенность модели SCC.
В основном – для обучения*

В обзоре [1] нами уже частично приводились данные о слабой распространенности модели “причинного пирога” в реальных эпидемиологических исследованиях. За прошедшее время ничего не изменилось.

Был осуществлен поиск в PubMed (февраль 2021 г.) на следующие точные сочетания:

“Sufficient causes model” – 5 публикаций (1993–1997);

“Sufficient-component cause model” – 22 публикации (2013–2020);

“Causal pie” – 11 публикаций (2010–2018).

Понятно, что три выборки пересекаются.

Для контроля:

“Hill criteria” – 257 публикаций (1991–2021);

“Hill’s criteria” – 104 публикации (1986–2021);

“Bradford Hill criteria” – 200 публикаций (1991–2021).

Можно согласиться с мнением Greenland S., Brumback B., 2002 [12], что модели SCC требуют характеристик механизмов для отдельных исследуемых единиц. Но такие данные имеются редко, и это объясняет, почему модели SCC мало используются, кроме обучающих примеров⁵⁸.

Видим, что именно так все и есть, а придумать “задним числом” иллюстративных примеров компонентных причин, как мы из радиационной эпидемиологии в [1] и здесь выше (рис. 2), а также иные авторы из медицины [27] (и др.), можно сколько угодно. Модель SCC, действительно, полезна для углубленного осознания мультифакторальной причинности событий (потому она и входит почти во все известные нам пособия по эпидемиологии; см. выше), но широко использовать ее на практике вряд ли возможно. Именно обучающая наглядность, на наш взгляд, обеспечивает живучесть модели SCC при ее столь слабом применении.

**Контрафактическая
(потенциальных исходов) модель**

“Например: “Если бы был сепсис, пациент бы умер” и “если бы не было сепсиса, пациент бы не умер”. Нам не нужно вдаваться в подробности этой широко известной теории, потому что сама ее основа слишком умозрительна и выходит за рамки любой доступной реальности”.

К. Sadegh-Zadeh. “...Philosophy of Medicine” (2015) [202]⁵⁹

Терминология и основные понятия

Оригинальное наименование модели: “Counterfactual model” или “Potential outcomes model”; это синонимы [12, 199, 241, 242].

Использование нами русскоязычного термина “контрафактический” (т.е. с “а”) здесь и ранее [1, 2, 5–7, 9], вместо имеющегося в орфографическом словаре “контрфактический” (без “а”), обосновано тем, что в отечественной философской литературе, посвященной подходу *от противного*, встречается только первое наименование (см. в [7, 9]).

Чтобы не плодить лишние обороты, мы, аналогично, постарались найти относительно весомые источники, в которых имелся бы перевод термина “Potential outcomes”. Поиск через Google дал немного. В переводе монографии по эконометрике [243] термин представлен в единственном числе – “потенциальный исход”; в других случаях (учебные программы и пр.; ссылки не приводятся) – “потенциальные исходы”.

Контрафактический подход (counterfactual approach; не модель, а философское понятие), равно как “Эксперимент контрафактический” (критерий Хилла), с примерами из радиационной эпидемиологии, подробно разобраны нами ранее [7, 9].

Важными являются понятия *контрафактического идеала* и *контрафактического контраста*. Теоретически это синонимы, но в реальной практике появляются замены/имитации второй категории. Термин (не понятие) “counterfactual ideal” впервые встретился нам, вкупе с “counterfactual contrast”, в “Modern Epidemiology” Rothman K.J. et al., 2008 [13]. Еще ранее “counterfactual contrast” был назван в 2002 г. [244].

Контрафактический идеал – это теоретический сценарий, когда индивидуум или группа людей, подвергавшихся воздействию, сравниваются с тем же самым индивидуумом или с той же самой группой, но без воздействия. Причем *в то же самое время* [24, 35, 40, 133, 154, 203, 244–248].

Поскольку подобный сценарий на деле невозможен, то, как указано в [244], используется подход по замене идеального “контрафактического контраста” на реально возможные группы сравнения, т.е. *контроли* [59]). Валидность подобной замещающей стадии (приближение к идеальному контролю) определяет валидность оценки эффекта [244]. А получение абсолютного знания о причинности для медико-биологических дисциплин невозможно в каком угодно эксперименте, повторим это вслед за D. Hume, A.V. Hill (подробно см. в [1, 7–9]) и J. Cornfield (см. выше эпиграф [122]).

Как приближение к контрафактическому идеалу рассматриваются исследования на однояйцевых близнецах [35, 248–250].

Таким образом, понятие “контрафактический” имеет два значения [7, 9]:

- Контроль, группа сравнения (попытка приближения к “контрафактическому идеалу”).
- Методический подход (или дизайн) по устранению воздействия (реальный или воображаемый) с последующим наблюдением за эффектом или его теоретической оценкой.

На последнем подходе и построена *контрафактическая модель* или *модель потенциальных исходов* для установления причинности [12, 13, 24, 199, 203, 206, 207, 241, 242, 247–252].

Исторические вехи

Этот момент также разбирался нами ранее, но – для контрафактического *подхода* или условия *в целом*, единого для философии, гносеологии, истории, статистики и эпидемиологии [7], а не для практической формализованной *модели* установления эпидемиологической причинности. Во многих источниках [13, 23, 49, 73, 198, 241, 253] отмечается, что впервые контрафактический подход для определения причины был предложен D. Hume в XVIII в. [254], причем замечено это оказалось спустя более 200 лет – в статье от 1973 г. философа D. Lewis (США) [214] (подробнее см. в [7]). До второй трети XX в. теория была разработана слабо, но подход, понятно, вполне использовался в названных выше дисциплинах (см. в [7, 13]).

Истоками собственно контрафактической *модели* причинности, с соответствующим формализмом, считается [241, 242, 248, 255] работа Neyman J. et al., 1923/1990 [256]. На ее основе D.V. Rubin предложил в 1974 г. [257] и, затем, развивал вплоть до второй половины 2000-х гг. (см. в [13, 242, 247, 248, 253]) свою “Rubin’s Causal Model” или “Potential Outcomes Model” для рандомизированных и обсервационных исследований в области статистики, социологии и психологии. В эпидемиологии подобную формализованную разработку осуществили S. Greenland, J.M. Robins

и J. Pearl [139, 258, 259] (и мн. др.), начиная, вероятно, с 1986 г. [258]. Весомый вклад в теорию контрафактического подхода в эпидемиологию внесла также публикация Maldonado G., Greenland S., 2002 [244] (подробнее см. в [7]).

Сущность контрафактической (потенциальных исходов) модели

Что такое контрафактический подход [7, 9], на основе которого создана контрафактическая модель? Кратко он формулируется так: “Если бы *A* произошло, то и *B* бы произошло. Чтобы быть контрафактическим, т.е. “против факта”, это *A* должно быть ложным или несоответствующим нашему миру” [253]. Отсюда и “альтернативные миры” D. Lewis (1973), в которых кенгуру могли бы существовать, не имея хвостов [7]. Понятно, что, вследствие недостижимости в качестве контроля “контрафактического идеала”, последний и является объектом таких ненаблюдаемых, фантастических миров.

Контрафактические контроль, подход и модель, собственно говоря, это весь мир экспериментальных и обсервационных дисциплин. Где выделенный или сформированный контроль – там и подход с моделью. Где оценка двух исходов, без воздействия и с таковым – там и контрафактическое. Как во втором романе про “Ходжу Насреддина”: “Кругом – волшебная трава! Простой даже и нет, все – волшебная”. Контрафактическая. Недаром автор объемного пособия по философии медицины обратил внимание на обыденность и тривиальность данной теории, вкупе с ее нереальностью, что было представлено в эпиграфе выше [202]. Ибо наблюдать истинный контрафактический мир и подобные ситуации невозможно – все контроли, как сказано, носят приближенный, суррогатный характер.

Тем не менее – эксперимент с контролем и опытом *in vitro*, на животных и в RCT, равно как “контрольная” и “опытная” группа или “наблюдаемое” и “ожидаемое” (observed/expected) число инцидентов/смертей в эпидемиологии – это все контрафактические модели, хотя и не слишком формализованные.

Некоторые исследователи полагают, что тот, кто использует в эпидемиологии даже тривиальные статистические модели и т.п., обязан понимать и называть эти модели [260], в то время как другие авторы считают это “медвежьей услугой”, заставляющей поверить в то, что если кто-то использует сложную модель для обсервационных данных, то он обязательно выявит истинную причинную связь, что не так [261]. То есть получается нечто вроде описанного Мольером “разговора прозой”.

Формульное представление указанной модели может быть разным и зависеть от поставленных задач [242, 247–252]. В качестве примера приведем построение модели потенциальных исходов для воздействия из социологической работы [242].

Участие в программе называется “терапией” (T , где $T=1$, если терапия получена, и $T=0$ в противном случае). В модели учитываются единицы (участники программы) и показатель результата Y . До начала терапии у всех единиц есть два пока еще нереализованных результата, “потенциальных исхода”: $Y(1)$ – потенциальный результат человека, подвергавшегося терапии, а $Y(0)$ – потенциальный результат того же индивидуума, не подвергавшегося лечению. В идеальном мире участники программы испытают оба результата, и эффект терапии будет измеряться как разница между этими двумя (потенциальными) результатами [$Y(1) - Y(0)$] для каждого индивидуума. Средний эффект терапии будет равен подобной разнице для всех участников. Но это, как сказано, невозможно, поскольку получение потенциального результата $Y(1)$ означает, что потенциальный результат $Y(0)$ не может наблюдаться ни для одного конкретного человека, и наоборот. *Наблюдаемый* результат (Y_{obs}) равен $Y(1)$ для лиц, отнесенных к экспериментальной группе, и равен $Y(0)$ для лиц, отнесенных к контрольной группе, т.е. $Y_{\text{obs}} = T \times Y(1) + (1 - T) \times Y(0)$.

Контрафактическая модель представляет оценку результатов как проблему отсутствия данных: $Y(0)$ отсутствует для лиц в экспериментальной группе, а $Y(1)$ – в контрольной группе. В реальных экспериментах делают выводы об этих пропущенных значениях, предполагая, что при рандомизации (или при случайном отборе пар в обсервационных исследованиях) группы эквивалентны, и, таким образом, средний наблюдаемый результат в контрольной группе эквивалентен (ненаблюдаемому и нереализованному) среднему потенциальному результату в группе терапии. Точно так же делается вывод, что средний наблюдаемый результат в группе терапии эквивалентен (ненаблюдаемому и нереализованному) потенциальному среднему результату в контрольной группе. И т.п., далее развивается формульный аппарат для вывода о вероятности исходов. Но выводы об эквивалентности наблюдаемого и потенциального исхода верны только в том случае, если экспериментальная и контрольная группы действительно эквивалентны. В RCT случайное распределение единиц увеличивает вероятность достижения эквивалентности. В обсервационных исследованиях статистические подходы, основанные на гипотетической модели, могут использоваться для выявления и контроля неэквивалентности групп, а также для последовательной оценки эффекта терапии [242].

Важным является то, что спрогнозированные исходы для каждого индивидуума (а их может быть много, учитывая таблицу 2×2 и аналогичные, более сложные таблицы [242, 247, 251]) остаются потенциальными до проведения воздействия (или воздействий) [13].

Модель позволяет оценить влияние смещений (bias) [13, 24, 247, 251] и конфаундеров [12, 13, 24, 241, 258].

Согласно [12], контрафактическая модель не является детерминированной, поскольку потенциальные исходы могут быть параметрами вероятностных распределений (например, ожидаемого возраста на момент смерти), а не непосредственно наблюдаемыми событиями (например, фактического возраста на момент смерти). Другие авторы отмечают либо детерминизм [133], либо “условный детерминизм” контрафактической модели [228].

Называется интеграция контрафактической модели с моделью SCC [207, 262] и с рассмотренной далее моделью прямых причинных графов (DAG) [206, 251]. В работе [207] авторы считают модель SCC “более тонкой”, чем контрафактическая.

Критика контрафактической (потенциальных исходов) модели

Само наименование “альтернативные миры” [7] звучит фантастически. Как безапелляционно указано в пособии по философии медицины от 2015 г. [202]:

“Абсолютно невозможно обосновать их в мире опыта и проверить эмпирически. Поэтому сторонник контрафактического утверждения не сможет разумно ответить на основной эпистемологический вопрос “Откуда вы это знаете?”. Несмотря на их интригующую метафизическую ауру, нельзя полагаться на контрафакты и плодотворно их использовать в эмпирически-практических областях, таких как медицина. Принятие диагностико-терапевтических решений не должно основываться на непроверяемых этиологических предположениях, если доступны более надежные знания. Кроме того, было убедительно продемонстрировано, что контрафакты приводят к абсурдным выводам, и поэтому их следует избегать в научных рассуждениях”⁶⁰.

Сходные послышки имеются и у других авторов [259, 263]: при рассмотрении причин прошлых событий представляются такие ситуации, которые никогда не встречались и, таким образом, не могут быть наблюдаемы и эмпирически тестируемы. В результате причинные выводы, основанные на контрафактическом подходе, зависят как бы от непроверяемых допущений. Это же самое

сказано в энциклопедии по статистике [253] со ссылками на [263, 264].

Понятно, что подобная критика касается только одного аспекта: контрафактических конструкций и фантазий, наподобие упоминавшихся нами ранее [7] из истории (“Гражданская война в США... контрафактическая история...”) или гносеологии (типа поговорки “Если, да кабы...”). Ни сам контрафактический подход, ни эксперимент, ни контроли в исследованиях никакая критика затронуть не может.

*Контрафактический подход, эксперимент
и модель в радиационных дисциплинах*

Нами уже указывалось [7], что для экспериментальных дисциплин радиационного профиля (радиобиология, радиационная генетика и т.д.) контрафактический *подход* обычно предусмотрен априори – в виде формирования контрольной группы. Для радиационной экологии, радиационной гигиены и радиационной эпидемиологии нами в [7, 9] были представлены примеры контрафактического *эксперимента* (полужэксперимент – semi-experiment Хилла [11, 34]). Использование же формализованной контрафактической *модели* (потенциальных исходов) для радиационной эпидемиологии нами не было обнаружено (поиск в PubMed на “Counterfactual model&Radiation”). Исключение составили исследования по оценке предотвращения рака кожи путем иного сценария облучения УФ Солнца [265, 266] и контрафактическое моделирование частоты рака мозга в 1985–2014 гг., предусматривающее как бы отсутствие тогда мобильной связи [267].

Возможно, приведенная выше критика является оправданной именно для подобных исследований.

Что же касается контрафактического подхода для собственно радиационной эпидемиологии, исследующей эффекты ионизирующей радиации, то, по-видимому, все известные умоглядные примеры пока что собраны в наших обзорах [7, 9].

**Модель направленных ациклических графов
или каузальных диаграмм**

“Рисунок 4 [перечеркнутый список критериев Хилла] был представлен во введении к выводу о причинности в 2018 г. [268], с однозначным сообщением, что как только вы сможете нарисовать направленный ациклический граф, вам больше не понадобится Хилл и его печальный, устаревший (sad, time-expired) эмпиризм”.

Davey Smith G., 2019 [142]

Терминология

Оригинальное наименование модели: “Directed acyclic graphs (DAGs)” или “Causal diagrams”; синонимы [13, 14, 27, 56, 59, 203, 251, 269–271]. Иногда встречается “Directed *and* acyclic graphs” [203]. Согласно PubMed, впервые MeSH “Directed acyclic graphs” встречается в статье 1984 г., посвященной компьютерным моделям в физиологии [272].

В русскоязычных источниках присутствуют наименования “направленный” и “ориентированный” графы (см. в Интернете). В отечественных и украинских публикациях используется первое название, но соответствующие дисциплины в большинстве не являются медико-биологическими – это информатика, искусственный интеллект, обучающие модели и т.п. [273–276] (в качестве обзора см. [276]). Все такие применения DAG в работах российских и украинских авторов лежат в рамках Байесовской сети [273–276], включая единичные (судя по PubMed и списку литературы в [277]) применения Байесовских сетей в российских исследованиях в медико-биологических областях [277–279].

*Принципиальная суть модели DAG – визуализация
детерминированной или вероятностной
причинности. Байесовские сети*

Как сказано в работе Золотин А.А., 2017 [276], образ мышления и запоминания, свойственный живым существам, базируется на причинно-следственной связи, и рассуждения можно схематично отобразить так: объект X влияет на объект Y ; порой с помощью объекта Z . Отсюда и конструирование интуитивно понятной и образно-иллюстративной модели, которая как бы проецирует образ мышления путем визуализации зависимостей между элементами (или объектами), соответствующими некоторым утверждениям, с помощью направленного пути в графе. Вершинами последнего являются эти элементы-утверждения (узлы – “nodes”), а ребра, арки или связи (“edges” [59, 280] или “paths” [59, 281]; “arcs” или “links” [280]) представляют причинно-следственные зависимости между ними:

$$X \rightarrow Y$$

$$X \rightarrow Z \rightarrow Y$$

Значение X на первой схеме является “родительским” узлом (“parent” [59, 281], “ancestor” [13, 59]; русскоязычный термин взят из [274, 277]), а Y – узлом-“потомком” (“child” [59, 280], “descendant” [13, 59]; русскоязычный термин взят из [277]). На второй схеме, понятно, Z является одновременно и тем, и другим, а Z и Y – оба являются “потомками” X , а X и Z оба являются предками (ancestors) Y [282].

“Ациклическим” граф называется, поскольку он не включает направленный путь, начинающийся и заканчивающийся в одной и той же вершине [13, 203, 274, 280–284]. То есть DAG содержит одну направленную стрелку, которая соединяет узлы в структуре графа, ведь поток информации течет из “прошлого” в “будущее”, по направлению стрелок. Никакие пути не возвращаются к родительскому узлу: они направлены от причинной переменной к переменной эффекта [281]; никакая переменная не является ее собственным родителем или собственным потомком [280]. Если стрелки направлены в обратном направлении, то такие пути называются “беспричинными” [282].

Теория графов [12–14, 259, 281] в той или иной степени используется в самых разных дисциплинах, включая медицину [277–279] и эпидемиологию [12, 13, 280, 281].

Причинная зависимость, как было видно выше и в [1], может носить как детерминированный, так и вероятностный характер. Для первого случая нами не было обнаружено построений или примеров DAG ни для какой дисциплины, хотя это, понятно, не значит, что таковых нет. Но ясно, что вполне можно построить граф для, например, тканевых (детерминированных) эффектов радиации [218] любой сложности, скажем так:

Локальное облучение в дозе 10 Гр →

→ Ожог ← Нахождение в очаге пожара

Вокруг этих узлов и путей можно сформировать не один промежуточный путь, также имеющий детерминированную причинность: отсутствие защиты от излучения или огня, экстремальная ситуация и пр.

S. Greenland и J. Pearl в 2011 г. [14] называли три уровня интерпретации графических отношений в DAG: вероятностный, причинный и функциональный. Основное применение ациклических графов предусматривает вероятностную причинность, когда графы входят [271, 273–279, 283, 284] или не входят [13, 14, 206, 251, 259, 270, 271, 281, 284–286] в Байесовские сети (использование DAG в вероятностной теории имеет отличие от использования в эпидемиологии [284]).

Байесовские сети (Bayesian networks), или “Байесовские сети доверия” [273–276, 287], используются для моделирования областей, характеризующихся неопределенностью, обусловленной: недостаточным пониманием предметной области, неполным знанием ее состояния в момент принятия решения, случайным характером механизмов, определяющих поведение этой области, или комбинацией названных факторов [274]. Байесовская сеть представляет собой вероятностно-графическую модель, которая отображает вероятностные зависимости множества пе-

ременных и позволяет делать вероятностный вывод с помощью этих переменных [276, 277]. Сеть включает, таким образом, соответствующий DAG и набор вероятностных распределений, определяющих силу отношений зависимости, закодированных в графической структуре — таблицу условных вероятностей узлов [274, 276, 277, 288]. Эта таблица рассчитывается с использованием существующих наборов интересующих данных (например, клинических или эпидемиологических), иных данных, полученных из опубликованных исследований и/или из мнений экспертов. Состояние предыдущего или родительского узла будет влиять на вероятность состояния дочернего узла [288].

Согласно [276], название “байесовские” связано не с байесовскими методами, а скорее, с байесовским правилом вероятностного вывода. Далее останавливаться на деталях байесовских сетей не представляется целесообразным. Хотя, судя по PubMed, этот подход в последние десятилетия широко используются в медико-биологических дисциплинах: на точное сочетание “Bayesian network” обнаруживается 2856 ссылок (1992–2021 гг.; дата поиска 05.03.2021). Но только три из них оказались российскими [277–279].

Малая распространенность модели DAG в эпидемиологии

В имеющихся у нас отечественных пособиях по эпидемиологии никаких упоминаний о графах и каузальных диаграммах не обнаружено (шесть пособий за 2006–2013 гг.; основные ссылки представлены ранее [2, 5, 7, 9, 10]). В пособии Власов В.В., 2006 [44] имеется раздел по байесовскому подходу, а в пособии под редакцией В.И. Покровского и Н.И. Брико от 2012 г. (ссылку см. в [9]) есть одно упоминание о теореме Байеса. Это все. Выходит так, что и байесовские сети в отечественных медико-биологических дисциплинах практически не используются (известные исключения для медицины были названы [277–279]), в отличие, как указано выше, от ситуации за рубежом.

Аналогичным образом, в отечественных и переводных пособиях и обзорах по ЕВМ и клиническим исследованиям ничего не найдено ни про DAG, ни про Байеса (Гринхальх Т., 2006; Мелихов О.Г., 2013 (см. в [9]), Реброва О.Ю., 2001; Спасов А.А., Черников М.В., 2007; Оганов Р.Г. (редактор), 2010 и др.; ссылки не приводятся). В западных пособиях по ЕВМ (например, Howick J., 2011; ссылку см. в [6, 9]) разработки на основе положений Байеса, однако, отображены.

Такая же ситуация с редким использованием теории графов, как сказано, и в отечественных медико-биологических публикациях. Между тем, согласно [289], модель DAG в эпидемиологии на-

считывает уже около 50 лет [290], а если взять ее популяризацию – то более 20–30 лет [269, 276] (см. также следующий подраздел).

Впрочем, сведения о модели DAG и о байесовских сетях отсутствуют и во многих западных учебниках по эпидемиологии: они не найдены нами в 28 пособиях (2004–2019) и найдены в 12 пособиях, порой только в последних изданиях (2000 г. + 2011–2020 гг.). Среди этих двенадцати всего в четырех рассматриваются теорема и вероятности Байеса (но два из них посвящены социальной эпидемиологии, а одно – статистике и каузальности в эмпирических дисциплинах в целом). Таким образом, в 70% обучающих западных источников по теме каузальная модель DAG не рассматривается, несмотря на отраженные в эпиграфе к настоящему разделу пафосные рассуждения. Обратная картина – с уже “перечеркнутыми” [142, 268] критериями Хилла, данные о чем мы надеемся представить позднее.

Тем не менее поиск в PubMed на точное сочетание “directed acyclic graph” выдает 448 результатов (1984–2021); 60 из них отвечают конструкции [“directed acyclic graph” and causation] (2007–2021 гг.; дата поиска 05.03.2021). В систематическом обзоре Tennant P.W.G. et al., 2020 [271] объемом 135 страниц, авторы поставили целью дать подборку всех исследований из области медицины и здравоохранения (health research), в которых бы использовали модель DAG. Коллекция включает 234 публикации (2003–2017); среднее число узлов-объектов на графах составило 12, а ребер-путей – 29 [271]. То есть, на наш взгляд, установленная причинная сеть в целом достаточно непростая.

Исторические вехи

Официально впервые DAG были изобретены эволюционным генетиком Sewall Green Wright (1889–1988; США) в начале 1920-х гг.; в 1923 г. им были сформулированы три правила, регулирующие пути в графах [281]. После генетики модель стала использоваться в иных областях, включая социологию [286], откуда, согласно [289], ее в 1973 г. ввел в эпидемиологию M. Susser [290], взяв из работы H. Blalock [291]. Таким образом, заключает A. Morabia в статье 2015 г. [289], M. Susser в одиночку пошел по пути, который много лет спустя проделали иные эпидемиологи.

Разработку вероятностных моделей на основе теории DAG, как сказано в [276], в информатике связывают с именем Judea Pearl (род. 1936; Израиль – США), с конца 1960-х гг. проводившего исследования в области искусственного интеллекта и т.п. В монографии 1988 г. J. Pearl обобщил свои изыскания, предложив новый подход к построению вероятностных моделей, основанный на DAG [292]. Позже автор продолжил свои разра-

ботки для разных областей, включая эпидемиологию [14, 139, 259, 269, 283]; в последнем случае с J.M. Robins и S. Greenland [14, 259, 269]. Однако еще ранее, в 1980 г., при оценке риска кесарева сечения на основе мониторинга плода, S. Greenland и R. Neutra представили настоящий причинный DAG для определения эффекта конфаундеров [293]. А в 1987 г. J.M. Robins опубликовал работу по использованию графической модели для определения причинности смерти после различных воздействий [285].

Потому еще вопрос, у кого приоритет, включая упомянутого M. Susser [289, 290], в принципе – сторонника причинных критериев [41, 42, 144–146, 197, 290] (см. также в [10]).

Работ трех указанных авторов, J.M. Robins, S. Greenland и J. Pearl, на тему использования DAG для доказательства причинности в эпидемиологии весьма много, что видно даже по нашему списку [14, 139, 259, 269, 283, 285]. Это – основные идеологи данного подхода.

Далее материал по применению модели DAG в эпидемиологии рассматривается относительно кратко, хотя и со специальными примерами. На деле – это целый мир, отчасти эзотерический для большинства специалистов в биологии, медицине и эпидемиологии. И не только российских.

Аппаратная терминология модели DAG. Адаптация к эпидемиологии

Специальная терминология элементов и функций DAG представлена во многих публикациях, но – в разной полноте, от самой основной [13, 27, 203, 270, 281, 283, 284, 286, 288, 294] до детализированной [59, 282, 295]. Заметим, что две последние ссылки – это пособия по социальной эпидемиологии и социологии. Только в третьем издании (но не в первом от 2002 г. [177], а второго мы не имеем) по собственно эпидемиологии Bhopal R.S., 2016 [59] есть раздел с адаптацией словаря терминов DAG применительно к указанной дисциплине. Который мы и воспроизводим в табл. 4: во-первых, данный материал малодоступен и, во-вторых, он проясняет эпидемиологический смысл DAG. Была сделана попытка отыскать используемые в русскоязычных источниках параллельные термины (чтобы не придумывать свои “кустарные”), но то единичное, что найдено, не имело отношения к медико-биологической сфере.

Символами, используемыми в структуре DAG, являются [59]:

- Уже называвшиеся линии (ребро, арка, путь, связь [59, 280, 281]);
- Стрелка на линии показывает направление причинного эффекта;

Таблица 4. Терминология, используемая в аппарате DAG применительно к сходным понятиям в эпидемиологии (по [59]; дополнено)*

Table 4. Terminology used in the DAG apparatus in relation to similar concepts in epidemiology (according to [59]; supplemented)*

DAG terminology (русскоязычный аналог)	Epidemiology terminology for similar ideas	Перевод или пояснение
Ancestors (узел-предок [287] или прауродитель)	Distal causes	Дистальная причина (т.е. лежащая после). Все предшествующие объекты в причинной цепи, связанные с исследуемым воздействием
Back door pathway или backdoor (бэкторный путь)	Confounding variable(s) creating the association	Ложная ассоциация через конфаундер. Стрелки этого пути идут от конфаундера на экспозицию и на результат
Block/blocking (блокирование)	Eliminating an association through, for example, adjusting, stratifying, etc.	Устранение ассоциации посредством, например, корректировки, стратификации и других подходов по удалению конфаундеров и смещений
Blocked path (заблокированный путь)	An association that has been eliminated as it has been controlled for, e.g. by adjusting for confounders	Ассоциация, которая была устранена, например путем коррекции на конфаундеры
Blocking (блокировка ассоциации)	Presence of confounding or selection bias	Наличие конфаундера или смещения отбора
Child (узел-потомок [277])	Effect, outcome	Узел, отражающий эффект или конечное событие
Collapsibility (устойчивость ассоциации к конфаундерам)	The measure of the association is not affected whether examining stratified or overall, actual (crude) rates	Определение ассоциации не зависит от того, исследуются ли стратифицированные или общие, фактические (грубые) частоты
Collider (коллайдер [287])	A variable that is caused by both the exposure and outcome under study	Переменная, которая вызывается как воздействием, так и исследуемым результатом. То есть воздействие и результат оба являются причинами (родителями) коллайдера
Collider stratification bias или M-bias	Berkson's bias/selection bias. It arises from, for example, adjusting for a collider	Берксоновское смещение или смещение отбора. Появляется, если, к примеру, выполнить корректировку выборки на коллайдер
Conditioning (кондиционирование, обработка)	General term to include stratification, standardization, and adjustment in a model (conditioning means holding a variable constant)	Термин включает стратификацию, стандартизацию и корректировку модели (кондиционирование означает поддержание переменной в качестве постоянной)
Descendants (узлы-“потомки” [277])	Effects on potential causal path including mediators	Узлы отражают объекты потенциального причинного пути, включая медиаторы
D-connected	The postulated causal path is open (see open path and path)	Постулируемый открытый причинный путь (см. ниже)
D-separation (или d-separation [284, 294]) или D-unconnected (directional separation). Удаление направления)	The postulated causal path is closed	Постулируемый закрытый причинный путь

Таблица 4. Окончание

DAG terminology (русскоязычный аналог)	Epidemiology terminology for similar ideas	Перевод или пояснение
Identification (идентификация)	Analysis of associations to separate error/bias/confounding from causal effects	Анализ ассоциации на предмет устранения ошибки/смещения/конфаундера для причинного эффекта
Mediator [13, 14, 270, 271, 281, 286] Модератор [271, 285]	— —	Переменная, опосредующая эффект Переменная, влияющая на направление и/или силу связи между иными переменными
Nodes (узлы [274, 276, 277, 287])	Variables	Варианты, объекты, элементы причинной связи
Open path (открытый путь)	Potential causal relationship, i.e. association	Потенциальная причинная связь
Parent (родитель [274, 277])	Proximal cause	Проксимальная причина (т.е. лежащая перед)
Path (путь)	The route to potential causality, i.e. from A to B	Потенциально причинное направление
Vertices (вершины)	Variables	Варианты, объекты, элементы причинной связи

* Если не указаны ссылки, перевод наш (А.К.).

• Переменные могут быть выполнены в виде кругов, квадратов и пр. геометрических фигур, внутри которых вписан определяющий вариант символ;

• Сплошные или пунктирные линии со стрелками, расходящиеся от одного узла — конфаундер;

• Сплошные или пунктирные линии со стрелками, сходящимися к единому узлу, — коллаيدر;

• Пунктирная линия — недоказанная ассоциация;

• Отсутствие линий — сильное предположение об отсутствии прямого причинного эффекта между двумя переменными для каждого члена популяции (“сильная нулевая” гипотеза отсутствия эффекта) [282].

Основная цель модели DAG — устранение эффекта конфаундеров и смещения отбора

Как можно было видеть в том числе из цикла наших сообщений и сопутствующих им публикаций, при установлении истинной причинности ассоциаций необходимо устранить случайность, конфаундеры, смещения и обратную причинность [1–10].

Задача определения и элиминации конфаундеров ставилась с начала внедрения DAG как в социологические исследования [280], так и в эпидемиологию [258, 269, 285, 293]. Она остается таковой и ныне [14, 27, 56, 59, 203, 205–207, 241, 251, 270, 271, 281, 283, 284, 286, 294, 295]. Другая задача —

выявить коллаидеры (см. в табл. 4), чтобы определить те переменные, которые не являются конфаундерами и, потому, попытка их корректировки путем рандомизации, стандартизации, рестрикции, стратификации, мэччирования, иного урегулирования (adjustment) или мультивариантной техники [13, 27, 39, 40, 57–60, 94, 95, 98, 154, 246] приведет только к смещению отбора (selection bias) [14, 56, 59, 206, 251, 270, 271, 281, 284, 286, 295].

Мы не обнаружили примеров идентификации обратной причинности (reverse causation [5, 8]) с помощью DAG. Вероятно потому, что тогда граф уже не будет ациклическим. Сходные с нашими рассуждения имеются на западном сайте по машинному обучению, где некто попытался представить схемы DAG с обратной причинностью из медицины, но был опровергнут именно названным аргументом [296].

Далее мы не станем углубляться в тонкости применения модели DAG в медико-биологических дисциплинах. Примеры графов, приводимые в некоторых западных пособиях по эпидемиологии либо совсем просты (как схемы кадета Биглера), либо не всегда адекватны по интерпретации нарисованного, хотя, конечно, это только наше мнение, не слишком компетентное применительно к проблеме. Тем не менее и мы позволим себе, в виде иллюстрации основных задач указанной причинной модели, собственные построения из области радиационной эпидемиологии. Тоже, конечно, простые (рис. 3).

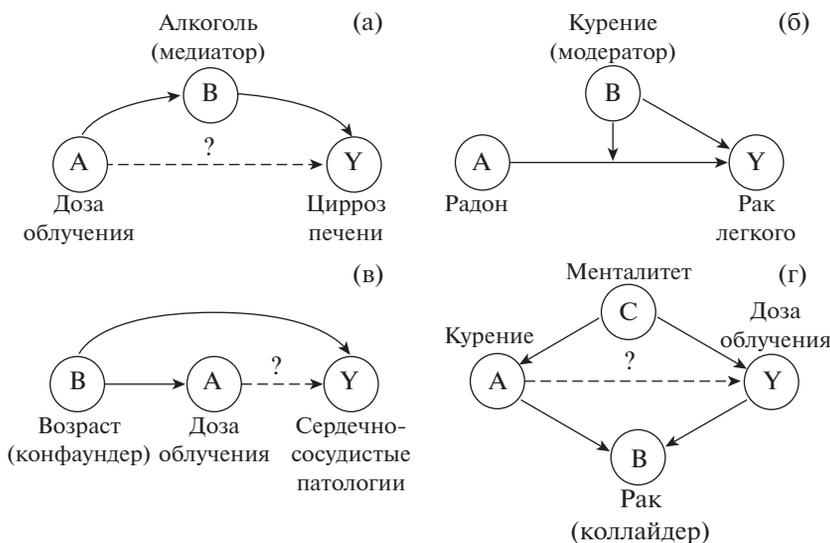


Рис. 3. Модели DAG для оценки причинности ассоциаций в радиационной эпидемиологии: а – эффект медиатора (алкоголя) у работников ядерной индустрии Франции [223]; б – эффект модератора (курения) при индукции радоном рака легкого [65, 118, 157, 161, 221]; в – эффект конфаундера (возраста) на связь между дозой облучения и частотой сердечно-сосудистых патологий [218]; г – конфаундер (менталитет) и коллайдер (частота раков) при оценке ассоциации между дозой облучения и частотой канцерогенеза у работников ядерной индустрии Японии [222].

Fig. 3. DAG models for assessing the causality of associations in radiation epidemiology: а – mediator effect (alcohol) in workers in the French nuclear industry [223]; б – the effect of the moderator (smoking) upon induction of lung cancer by radon [65, 118, 157, 161, 221]; в – the effect of confounder (age) on the relationship between the radiation dose and the incidence of cardiovascular pathologies [218]; г – confounder (mentality) and collider (cancer rate) in assessing the association between the radiation dose and the carcinogenesis rate in workers in the nuclear industry in Japan [222].

Из DAG на рис. 3, а видно медирующее действие алкоголя на индукцию цирроза печени у работников ядерной индустрии Франции [223] (этот пример уже приводился выше). Радиация в тех дозах, которые могут получать работники, не вызывает поражений печени [218], но установлено, что уровень потребления алкоголя оказался связанным с дозой радиационного воздействия [223]. Объяснений данному факту может быть два: прямое (длительность занятости отражается на обеих нагрузках, и лучевой, и алкогольной), и опосредованное авантюрным менталитетом, который может, вероятно, способствовать и потреблению алкоголя, и согласию на более опасные работы. Но с формальных позиций именно алкоголь (при отсутствии гепатита и химических воздействий) будет медиатором цирроза печени.

DAG на рис. 3, б отражает известный факт потенцирования курением (модератор) эффекта радона на частоту рака легкого [65, 118, 157, 161, 221].

При непродуманном дизайне исследования эффект конфаундера (возраста) может имитировать связь между дозой облучения и частотой сердечно-сосудистых патологий как в ядерной индустрии, так и в области радиотерапии. Это иллюстрирует DAG на рис. 3, в. Бэкдорный (backdoor; обратный; см. табл. 4) путь, который идет от дозы облучения через возраст к сердечно-сосудистым патологиям, может быть заблокирован

(block/blocking) путем урегулирования (conditioning) на возраст. То есть – с помощью названных выше стандартных эпидемиологических процедур, устраняющих конфаундер.

DAG, однако, не позволяет оценить качество или силу эффекта конфаундера; он только идентифицирует последний [284].

Наконец, на рис. 3, г приведен наиболее сложный случай, соответствующий также ранее рассмотренной ситуации с работниками ядерной индустрии Японии [222]. Интенсивность (тяжесть) курения оказывается связанной с дозой облучения, каковая связь не имеет биологического механизма. Но налицо вмешивающийся эффект конфаундера – менталитета, который, как уже говорилось, способен определять и склонность к вредным привычкам (курение, алкоголь), и снижение внимания к своему здоровью, и более частое согласие на опасные работы, и меньшее следование нормам радиационной безопасности. Если заблокировать бэкдорный путь от дозы радиации к курению, скажем, стратификацией по менталитету (хотя бы по частоте прохождения медосмотров [222] и, скажем, регулированием на фактор алкоголя; останутся просто курящие, без “легкомыслия”), то кажущаяся ассоциация “курение → доза радиации”, скорее всего, будет заблокирована.

Из DAG на рис. 3, г виден также коллаيدر: это частота раков. Данную частоту могут увеличивать и курение, и радиация. Конечно, курение влияет на названный показатель намного сильнее, чем реально накопленные работниками ядерной индустрии дозы облучения [118, 219, 220, 297]. Но если представить себе, скажем, плутониевое и тому подобные производства 1940-х — начала 1950-х годов в России, Великобритании и США [297], то вывод будет совсем не так однозначен.

Допустим, кто-то желает прояснить связь между курением и накопленной дозой радиации путем анализа некоего промежуточного элемента — частоты раков. Но обратный путь от дозы радиации к курению через частоту раков — это не бэкдорный путь; он заблокирован коллаидером. Поэтому если пытаться стратифицировать или еще как-то стандартизировать частоту раков (например, выделив только раки легкого, только лучевые раки и т.д.), то устранить обратный путь не получится. Результатом окажется лишь смещение отбора (selection bias) для раков. И выявить связь между курением и дозой радиации путем изучения частоты раков — не получится.

Частные корреляции и частный DAG (partial DAG)

Поскольку одна из главных целей модели DAG, как указано выше, это устранение эффекта конфаундеров, то ее нередко применяют для оценки частных корреляций.

Теория частных корреляций, разработанная в 1907 г. G.U. Yule [298], дает возможность выявить истинные корреляционные связи между двумя переменными в тех случаях, когда с ними попарно ассоциированы еще и иные факторы. К примеру, две варианты могут коррелировать друг с другом не только за счет истинной причинности, но также из-за того, что они обе меняются под влиянием некой третьей переменной. Поэтому если корреляция между двумя вариантами уменьшается при фиксировании третьей переменной, то это свидетельствует, что зависимость частично обусловлена последней. Иной раз, при устранении влияния “третьего фактора”, частная корреляция становится равной нулю, что указывает на полную обусловленность эффекта вмешивающимся фактором [299, 300].

Несмотря на тесную связь с понятием о конфаундерах, теория частных корреляций никак не отражена в известных нам десятках пособий по эпидемиологии, доказательной медицине и медицинской статистике — отсутствует даже термин. Это, конечно, пробел. Указанная теория излагается только в некоторых руководствах по специальному статистическому анализу, давая основу теории множественной регрессии [300]. Поскольку коэффициент корреляции Пирсона

представляет собой меру линейной связи между двумя переменными, то коэффициент частной корреляции — это мера линейной зависимости, которая все еще существует, когда линейные влияния одной или нескольких иных переменных устранены [300]. Используемая нами ниже стандартная формула расчета частной корреляции между двумя переменными, с учетом попарных корреляций с третьей, приведена во многих учебных руководствах [299].

Примеры частных корреляций можно найти в публикациях из различных дисциплин, но преимущественно — в рамках социологии, психологии, психиатрии (например, [301, 302]; всего обнаружено 9 работ) и геномики [303–307], в том числе для облученных клеток и организмов [308, 309]. Среди дисциплин радиационного профиля, согласно поиску в PubMed, найдено немного исследований, в которых применялись оценки частных корреляций:

- Определение связи между концентрацией радона в помещениях и риском лейкозов с учетом фона от γ -излучения, социоэкономического статуса и географических факторов [310];
- Множественный регрессионный анализ вклада различных факторов (включая облучение), в “эффект здорового работника” [311];
- Определение риска солидного рака в японской когорте пострадавших от атомных бомбардировок в зависимости от воздействия нейтронов и γ -излучения [312];
- Оценка частоты транслокаций после компьютерной томографии в зависимости от эффекта разных факторов [313].

Есть и другие единичные примеры, из области радиационной экологии [314], радиационной иммунологии [315] и т.п. [316]. Но в целом следует сказать, что в радиационной эпидемиологии, несмотря на нередко многофакторную причинность, теория частных корреляций практически не применяется. Дело, вероятно, в том, что в эпидемиологии, в отличие от таких дисциплин, как социология, психология и т.п., при оценке эффекта используются показатели риска, но никак не корреляции, которые для эпидемиологии имеют мало каузальной ценности [4].

Отображение частных корреляций в partial DAG (p-DAG). Модель DAG, графически отображающая связи между переменными, хорошо подходит для иллюстрации множественных ассоциаций с последующим выводом о необходимости расчета частной корреляции. Такие графы называют “partial DAG” или “p-DAG” [301, 317]. Этот термин в точном виде (полный и сокращенный) встречается нечасто: через PubMed выявлено всего два указанных источника. Но отображение в виде DAG и причинных сетей многофакторных зависимостей для оценки частных корреляций

относительно распространено в разных дисциплинах [302–304, 306, 307] (и др.), причем сети, порой, выглядят чрезвычайно сложными и поддающимися только компьютерному анализу [302, 303, 307].

На наш взгляд, подобная сложность далека от реальной практики установления причинности, ибо всегда будет слишком много флуктуаций в рамках неопределенностей. Тем не менее мы попытались представить свой, гипотетический пример расчета частной корреляции с использованием р-DAG из радиационной эпидемиологии. В качестве иллюстрации предлагается оценка истинной величины коррелятивной ассоциации между накопленной дозой облучения и риском смертности от солидных раков для работников ядерной промышленности. Используемые источники для этой категории работников в основном касаются Великобритании, но некоторые данные пришлось взять из исследований российских групп. Следует отметить, что нижеприведенный пример – только иллюстративный; он если и претендует на какие-то выводы, то лишь качественного характера. Пример основан на следующих пунктах, связанных с конкретными исследованиями работников ядерной промышленности и иных групп.

1. Для лучевого воздействия как такового известен факт зависимости между частотой смертности от рака и накопленной дозой радиации [65, 118, 157, 161]. Нами была предпринята попытка определить величину корреляции Пирсона (программа Statistica, ver. 10) для данной ассоциации применительно к работникам ядерной промышленности Великобритании. Использовались данные 6 исследований (1988–2014). Однако в тех случаях, когда сообщалось о кумулятивных экспозициях в пределах малых доз (до 100 мГр излучения с низкой ЛПЭ [65, 118, 161]), никаких ощутимых корреляций получено не было ($r = 0.210$; $p = 0.734$ [318]; $r = 0.283$; $p = 0.645$ [319]). В двух случаях, даже для диапазонов доз до 200 мЗв [320] и 400 мЗв [321], выявились вовсе отрицательные корреляции с дозой.

В оставшихся двух исследованиях, где были приведены данные по смертности от рака у работников, накопивших дозы до 520 мЗв [322] и 582 мЗв [323], нужная корреляция, наконец, появилась, и в одном случае – статистически значимая ($r = 0.591$; $p = 0.094$ [323]; $r = 0.834$; $p = 0.020$ [322]).

Последняя величина была условно принята нами за уместную для последующих расчетов⁶¹.

2. Доза, накопленная работниками ядерной промышленности, зависит от временного периода, в который они начали осуществлять и осуществляли свою деятельность. Ведь, согласно [325], нормы радиационной безопасности сильно ужесточались от десятилетия к десятилетию, и допустимая

доза уменьшилась на два порядка за менее чем 70 лет (с 1560 мЗв/год в 1925 г. до 10–20 мЗв/год в 1990-х гг.; эти сведения подробно отражены в наших публикациях [6, 297]). Если, исходя из данных работы [326], построить корреляцию между величиной средней накопленной дозы для работников ядерной промышленности Великобритании и серединой представленных в [326] временных диапазонов начала их работы, то выявляется значимая отрицательная корреляция: $r = -0.882$; $p = 0.001$.

3. Смертность от рака (или ее вклад в общую смертность) как в генеральной популяции [327], так и для работников ядерной промышленности [327–329], также зависит от временного периода. Цитированные источники, однако, были обнаружены нами только для работников ПО “Маяк”; для множества западных исследований, как ни странно, нам не известно разделение канцерогенных эффектов по временным периодам. В результате для расчета корреляции “Временной период – риск смертности от рака” использовались данные после оцифровки (GetData Graph Digitizer, ver. 2.26.0.20) графика на Fig. 1, b из [327], и искомые значения составили: $r = -0.790$; $p < 0.001$.

Кроме того, на частоте смертности от рака отражается, естественно, возраст индивидуума [330–333]⁶² и возраст на момент воздействия [334], но мы эти данные в расчет не включали.

Полученные на трех приведенных стадиях корреляции отражены в виде р-DAG на рис. 4.

Расчет по известной формуле [299] частной корреляции между накопленной дозой и риском смертности от рака для работников ядерной промышленности показывает, что, сравнительно с нескорректированной величиной, значение коэффициента корреляции Пирсона уменьшается в 1.76 раза. Это обусловлено тем, что в разные временные периоды (конфаундер) частота смертности от рака варьирует во всей популяции, а не только у радиационных работников, обеспечивая в прежние десятилетия больший вклад злокачественных новообразований в общую смертность⁶³. Полученные данные снижают степень лучевой атрибутивности канцерогенных эффектов у указанного контингента.

Модель DAG следует рассматривать как инструмент, но не как независимую новую концепцию причинности

Это утверждение включено в объемное пособие по эпидемиологии от 2014 г. [284]. И можно согласиться, что DAG “обеспечивают вспомогательный подход к визуализации причинных связей между воздействием и результатами эпидемиологических исследований” [284].

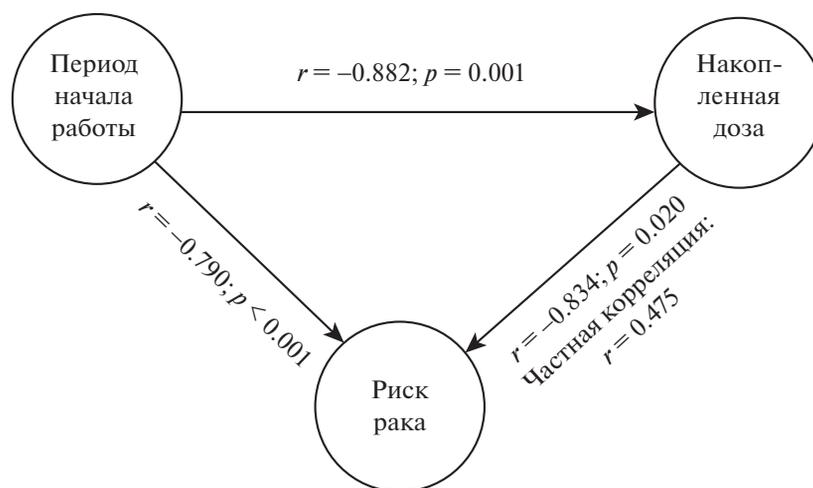


Рис. 4. Partial DAG (p-DAG) при расчете частной корреляции для зависимости между накопленной дозой облучения и риском смертности от рака у работников ядерной индустрии. Подробности в тексте.

Fig. 4. Partial DAG (p-DAG) at calculating the partial correlation for the relationship between cumulative radiation dose and cancer mortality risk in nuclear workers. Details in the text.

На наш взгляд, подобные модели – действительно только визуализация. Чтобы нарисовать каждую сплошную стрелку в DAG, необходимо, вероятно, проанализировать не один причинный критерий и использовать не один стандартный подход из классической эпидемиологии. Хотя нарисовать пунктирных линий из априорных соображений можно немеренно.

Уже после написания этих слов нам встретились аналогичные утверждения и других авторов. В [335] указано, что DAG могут быть полезны, но сами по себе они не должны сводить на нет причинно-следственные связи, а в Geneletti S. et al., 2011 [336], была сделана попытка совместить в каузальной методологии для геномики критерии Хилла и модель DAG. Отмечается следующее [336]:

“Мы используем DAG для визуализации сложных ассоциаций, но когда в нашем распоряжении есть только данные наблюдений, мы должны найти другие способы оценить: а) является ли конкретная ассоциация причинной, а не вызванной конфаундером или смещением (bias); б) каково направление этой ассоциации”.

И, далее:

“Один из способов ответить на этот вопрос – включить в схему с DAG предварительные знания, основанные на [критериях] Хилла (или аналогичных критериях), чтобы определить, какие грани-стрелки можно считать причинными”⁶⁴.

Иными словами, чтобы построить каждую стрелку в DAG, по мнению [336] необходимо использовать причинный подход, преимущественно основанный на все тех же критериях, что нами и было независимо сказано выше.

Чрезвычайная, порой, сложность DAG в медицине [277] (63 узла для дифференциальной диагностики гипертонии “при первичном обращении”), как и для рассмотренных нами в [1] “Сетей причинности” (“Web of Causation”), делает маловероятным какое-либо реальное применение. Это наглядные иллюстрации и, может, вспомогательные инструменты, но не методологии. Такие схемы красиво представлять во время учебного процесса.

Конечно, все это только наше мнение, хотя и подкрепленное, скажем, в [335, 336]. Формально же теория DAG на равных называется среди двух – четырех методов установления причинности, которые ныне рассматриваются вне причинных критериев [12, 13, 19, 24, 27, 40, 56, 59, 67, 139, 200, 202–204, 206, 207, 246, 282].

DAG в радиационных дисциплинах

Поиск через PubMed на [“directed acyclic graph” & radiation] дал пять источников, из которых три имели какое-то отношение к теме (в остальных “радиация” использовалась в эволюционном смысле – как “расхождение” и др.). Две работы, одной группы, имели отношение к радиотерапии [337, 338], а еще одна – к использованию байесовских сетей при оценке риска радиационной пневмонии [339]. Среди проанализированных в упоминавшемся систематическом обзоре [271] 234 работ по использованию DAG в медицине и здравоохранении, всего в одной публикации называлось облучение (радиотерапия) [340].

Таким образом, наши рис. 3 и 4 имеют, вероятно, некий приоритет.

Модель структурных уравнений

Это четвертая относительно известная модель установления каузальности вне критериев причинности, хотя она выделена и названа не во всех обобщающих тематических обзорах [12, 14, 19, 86, 205, 282]. Согласно [12], первоначально DAG были разработаны как интуитивное средство моделирования структурных уравнений, в котором сеть причинно-следственных связей моделируется системой формул и независимых предположений. Каждое уравнение показывает, как переменная индивидуального ответа (конечный результат, эффект, зависимость) изменяется при варьировании ее прямых (родительских) причинных переменных. Структурные уравнения с неизвестными параметрами выходят за рамки DAG, определяя функциональную форму эффектов, но не дают их точных значений. Таким образом, структурные уравнения являются алгебраическими, но не полностью количественными представлениями причинных отношений.

Здесь данная модель не рассматривается.

Метод количественной оценки ошибок (неопределенностей) эпидемиологического исследования С. V. Phillips

Наименование метода, точнее подхода, не выделено авторами в какую-то жесткую конструкцию, следующее: “Error quantification approach” (2004–2005) [16–18]. Метод представляет формальный интерес, поскольку С. V. Phillips и К. J. Goodman являются активными критиками подхода, основанного на критериях причинности, в частности Хилла [6, 9, 18, 33, 165]. Что же ими, как и К. J. Rothman выше, предлагалось взамен?

По сути ничего конкретного в плане установления именно причинности в эпидемиологическом и биологическом плане, и мы прошли бы мимо этой модели, если бы в статье о критериях Хилла от 2005 г., под авторством К. J. Goodman и С. V. Phillips, в энциклопедии по статистике в социальной психологии (“Behavioral Science”), в конце не было приведено следующее [18]:

“...обращение к набору критериев мало что дает для ответа на этот вопрос. Но недавние разработки в области методов количественной оценки неопределенности [17] создают инструменты для оценки вероятности того, что наблюдаемая связь обусловлена альтернативными объяснениями, которые включают случайную ошибку или смещение исследования (bias), а не причинную связь”.

Материал в работе [17] от 2003 г. показывает, что названное, по всей видимости, это разработка С. V. Phillips (хотя, порой, с разными соавторами), поскольку в [17] имеется много ссылок на соответствующие публикации, начиная с тезисов 1999 г.

Суть метода кратко заключается в том, что на каждом этапе эпидемиологического исследования необходима количественная оценка ошибки (метода, измерения и пр.; “неопределенности”) с последующим ее учетом. А не просто выдача конечного показателя с единственной, “точечной” ошибкой или доверительными интервалами. Оценку предлагалось проводить в том числе путем моделирования методом Монте-Карло и байесовского подхода. Иллюстрации начинаются с аналогии при округлении в математике значений до той или иной цифры после запятой с объяснениями, почему это имеет смысл в эпидемиологических исследованиях [17]. То есть так иллюстрируются количественные оценки неопределенности. Какое отношение данный подход имеет к нередко качественному определению эпидемиологических и биологических причинных зависимостей, понять не удается; публикаций на эту тему после 2005 г. [18] не обнаружено.

ЗАКЛЮЧЕНИЕ

В настоящем разделе ссылки, как правило, не приводятся — их можно найти выше.

Часть 2 сообщения 4 является составной по смыслу: в ней рассмотрены три отдельных направления, являющихся продолжением и, порой, завершением наших предыдущих исследований на тему критериев причинности [1–10].

Нельзя было пройти мимо попыток систематизации и построения иерархии (“рангов”) критериев причинности, выполненных на основе различных подходов. Здесь нами использовались две методологии: обращение к разработкам различных исследователей, т.е. как бы теория, и оценка путем определения частоты встречаемости того или иного пункта в тех работах, в которых критерии причинности использовались как способ установления каузальности эффектов. То есть как бы из практики.

У различных авторов и международных организаций иерархия критериев и частота использования конкретных пунктов варьируют. Но, если не считать априорно лидирующего критерия “Временная зависимость”, на первые два места следует поставить “Силу связи” и “Постоянство ассоциации” (последовательность их может быть и обратной) — сугубо индуктивные подходы. Третье место занимает “Эксперимент контрафактический” (т.е. “природный”), который, однако, доступен в эпидемиологии достаточно редко [7]. Более того, оценки в областях экологии и экотоксикологии продемонстрировали, что ни один из псевдо-экспериментальных дизайнов (до и после контрольных воздействий на окружающую среду) не является надежным [341] (в наших обзорах [7, 9, 10] на данную тему этот факт не приводился).

Поэтому на третье место может выходить критерий “Биологический градиент” (т.е. зависимость “доза–эффект”). В областях экологии и экотоксикологии, где не только эксперименты, но, порой, и обсервационные исследования на людях недоступны, ранги критериев специфичны: на первых местах находятся “Биологическое правдоподобие” и, вновь, “Эксперимент контрафактический”.

Другой важной темой является критика каузального подхода в обсервационных дисциплинах, основанного на критериях причинности. Было выявлено 10 направлений подобной критики, начиная от отсутствия доказательности самого индуктивного подхода, которому отвечает большинство критериев Хилла. Не вдаваясь в подробности описанного выше, скажем, что некоторые авторы (M. Susser, D.L. Weed) пытались дополнять список критериев некоторыми принципами, вводящими дедуктивные подходы, но все это представляется мало подходящим для практики эпидемиологии и здравоохранения.

Существует также мнение, что критерии причинности, являющиеся как бы “подтверждающими эффекты”, надо оставить только для принятия решений в области здравоохранения, поскольку с их помощью невозможно доказать каузальные связи в сфере “чистой науки”, где главное – фальсификация гипотез, т.е. “отрицание”. Вряд ли этот подход для эпидемиологии может быть принят всерьез.

Имелись и еще претензии к критериям, основанные на их “мозаичности”, мультидисциплинарности и субъективизме, равно как и неабсолютности (критерии “обременены оговорками и исключениями” [13, 52, 110, 138, 191]), но все подобные вопросы можно поставить почти перед любыми подходами в области естественнонаучных дисциплин.

Важным мог бы быть тезис о том, что доказательность подхода, основанного на причинных критериях, “никогда не была проверена экспериментально” [18, 20, 33], если бы не исследование Swaen G., van Amelsvoort L., 2009 [55], в котором, во-первых, была разработана методика рангового взвешивания критериев Хилла, и, во-вторых, проводилась ее проверка путем оценки вероятности канцерогенных эффектов для агентов, выбранных по списку МАИР (IARC) в категориях 1 и 2А. Частота согласованных (точных) прогнозов составила 82%; модель верно классифицировала 89% канцерогенов категории 1 и 71% канцерогенов категории 2А. Несмотря на возможность “круговой логики” (ведь МАИР нередко использует для подтверждения канцерогенности факторов именно критерии Хилла), тем не менее, некая проверка адекватности комплекса причинных критериев все же была проведена.

Наш цикл сообщений оказался бы неполным, если бы не были рассмотрены иные модели оценки причинности эффектов в эпидемиологии, вне каузальных критериев. В соответствующих обзорах обычно называют еще две, максимум четыре модели [12]. Первая и самая известная, которую называют порой идеологической основой современной эпидемиологии [19, 86, 142], это уже отчасти рассматривавшаяся нами ранее [1] модель достаточной компонентной причины K.J. Rothman (“пирог причинности”), подразумевающая детерминизм эпидемиологических эффектов – в отличие от позиции практической эпидемиологии и здравоохранения, в которой причинность определяется как вероятностная. Модель K.J. Rothman имеет значительные достоинства в описании и интерпретации причинности, которая всегда многокомпонентна и является результатом действия комплекса не одного фактора. Всего нами насчитано шесть пунктов, отражающих положительные стороны модели. Но восемь пунктов, нередко серьезных, отражают недостатки модели и ее во многом только теоретический и иллюстративный характер. Как сам K.J. Rothman в начале внедрения своей модели [215], так и другие авторы позже, приводили тезис о том, что она отражает как бы “чистую науку”, метафизический взгляд на причинность индивидуального уровня, в отличие от практики эпидемиологии с ее причинностью популяционного уровня.

Второй моделью установления каузальности является контрафактическая, также отчасти рассмотренная нами ранее [7]. Иное наименование ее – “модель потенциальных исходов”. По сути контрафактический подход, т.е. выстраивание альтернативных исходов, “от противного”, а также формирование контролей в экспериментальных и обсервационных дисциплинах, является обычной практикой, но “контрафактический эксперимент” в эпидемиологии и экологии, равно как в их приложениях для дисциплин радиационного профиля, как сказано, возможен редко [7, 341].

Третьей каузальной моделью является модель направленных ациклических графов или причинных диаграмм с общепринятой аббревиатурой “DAG”. Суть модели – в визуализации детерминированной или вероятностной причинности. Модель DAG входит также в байесовские сети. Распространенность DAG в эпидемиологии не слишком велика; основная ее цель состоит в устранении эффектов конфаундеров и смещений отбора. По мнению некоторых авторов [284], к которому мы присоединяемся, модель DAG следует рассматривать как инструмент, но не как независимую концепцию причинности.

Наконец, порой называется также модель структурных уравнений, которая, как и предыду-

щие модели, неспособна установить количественные причинные отношения и нами здесь подробно не рассматривалась. В процессе исследования был обнаружен также подход С.В. Phillips, заключающийся в попытках именно количественной оценки ошибок (неопределенностей) на всех этапах эпидемиологического исследования, который его авторы упоминают в контексте отрицания ими значимости критериев Хилла [18]. Но указанный подход не получил, видимо, развития (поскольку, судя по всему, тривиален) и не упоминается после 2005 г. [18].

Таким образом, в части 2 сообщения 4 нами были объективно рассмотрены, по-видимому, все известные, упоминающиеся и даже не упоминающиеся ныне модели установления причинности в эпидемиологии. Собраны все “за” и все “против” по каждой из них, и сомнительно, что по главным моментам можно добавить что-то еще. Наш анализ был полностью объективным, хотя и не беспристрастным, ибо, на наш взгляд, все приведенные вне причинных критериев модели и способы установления каузальности эффектов оказываются или качественными и иллюстративными, или представляют из себя лишь графический либо математический аппарат, служащий все тому же исходному подходу. Основанному на логике человеческого мышления и, если заглянуть вглубь, на все тех же причинных критериях.

Мы не раз упоминали ранее [1–10] и здесь, что международные организации из областей эпидемиологии, здравоохранения, экологии, экотоксикологии и пр., результатом деятельности которых является принятие решений и практические рекомендации, в своих официально сформулированных подходах используют именно причинные критерии, нередко называемые критериями именно Хилла. Широта их применения в самых разных направлениях впечатляет, в то время как использование остальных перечисленных моделей просто несравнимо.

Вопрос о применении критериев причинности (“критериев Хилла”) в самых разных обсервационных дисциплинах и самыми разными организациями (социология, экономика, юридическая область, психология, различные эпидемиологии, здравоохранение и прочие направления) мы надеемся рассмотреть в заключительной работе нашего цикла — части 3 настоящего сообщения 4.

ПРИМЕЧАНИЯ

1. “Here, then, are nine different viewpoints from all of which we should study association before we cry causation. What I do not believe — and this has been suggested — is that we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we accept cause and effect. None of my nine viewpoints can bring indisputable ev-

idence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question — is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?” [11] (здесь и далее — перевод А.К.).

2. В работе Phillips C.V., Goodman K.J., 2004 [33], под названием “Пропущенные уроки сэра А.В. Хилл”, сказано, что “Хилл советовал эпидемиологам избегать слишком большого акцента на значении статистического испытания, учитывая наблюдение, что систематическая ошибка часто является большей, чем случайная ошибка”. “Его предложения о том, как интуитивно (intuitively) оценивать причинную обусловленность, почти полностью теряются, если сводить все к контрольному списку (checklist)”. Таким образом, ведущий мировой медицинский статистик XX в. [2], А.В. Хилл относился к статистике, судя по всему, прагматично.

3. Некоторые авторы (например [13, 35, 37, 38]) рассматривали момент об абсолютности “Временной зависимости”, но никак не в связи с приведенной цитатой из Hill A.V., 1965 [11]. Только в пособии по эпидемиологии Aschengrau A., Seage G.R., III., 2014; 2020 [40] обратили внимание, сразу после цитирования из [11], на несуразность: “Мы согласны с сутью вывода Хилла, но мы полагаем, что одна часть его утверждения неверна — “Временная зависимость” является требованием причинности”. (“We agree with the essence of Hill’s conclusion; however, we think that one part of his statement is incorrect — temporality is a requirement for causality”.) Цитируем в связи с редкостью столь очевидного вывода, единичного для многих сотен источников. Наверное, это материал для психологии.

4. “While data to support some criteria may be incomplete or missing for any given assessment, these criteria offer a useful way to evaluate available information” [61].

5. “The strong consensus among epidemiologists is that conclusions about causation should not be drawn, if at all, until a number of criteria have been considered” [62].

6. “Today, the 1964 Surgeon General’s report and Hill’s subsequent causal criteria are routinely cited as authoritative statements of the proper method for assessing a body of etiological evidence” [63].

7. “It is generally accepted that the Bradford Hill criteria, as formulated in 1965, are still the most relevant criteria to be used in causal inference”. “It is generally accepted that the Bradford Hill criteria, as formulated in 1965, are still the most relevant criteria to be used in causal inference”. “Bradford Hill criteria still remain key components to causal inference” [55].

8. “We believe that Bradford Hill’s guidelines form a useful tool as they stand” [37].

9. “Various modifications of this list have been suggested, and many of the elements remain cornerstones of judgment on whether an exposure really does cause a disease, or whether an intervention is effective in preventing or treating disease” [56].

10. “Another proposed benefit of using Hill’s criteria was that these provided a framework within which allowance could be made for incomplete datasets and missing information” [38].

11. “Guidelines suggested by Bradford-Hill are often used to judge the confidence that a risk factor (e.g. radiation exposure) might cause increased incidence in a population of health outcomes (e.g. cancer)” [65].

12. “The Bradford Hill Criteria remain one of the most cited concepts in health research and are still upheld as valid tools for aiding causal inference” [66].

13. “...these have stood the test of time as a valuable way of assessing the strength of causal inference where an association has been demonstrated” [60].

14. “Notwithstanding the renewed interest in other models of causality over the past few years... Hill’s guidelines remain the cornerstone of causal inference for the practical epidemiologist and health policy expert” [27].

15. “In spite of the rapid development in theoretical epidemiology, they have remained at least as a reference point for causal thinking in review committees and for decision makers” [67].

16. “Hill also stressed that it was meaningless to combine criteria into scores because an association was not more likely to be causal if it met more criteria” [68].

17. “...he offered no means of deciding whether these aspects hold when considering a given association, no hierarchy of importance among them, and no method for assessing them to arrive at an inference of causation. Hill included the list of nine viewpoints in four editions of his textbook of medical statistics from 1971 through 1991..., without further elaboration than appeared in the original paper” [18].

18. Критерий “Прогностическая эффективность” (“Predictive performance”) был предложен М. Susser как уступка “Popperian Epidemiology” [10] и далее использован Г.А. Фох при формировании комплекса руководящих принципов причинности для экологии и экоэпидемиологии [89].

19. Факт наличия ассоциации выделен в критерий также М. Susser [10].

20. “The selection of criteria is more a matter of customary practice than theoretical rigor”, “...is more a matter of personal preference than of careful inquiry” [90].

21. В Gordis L., 2014 [39] сначала названы все критерии Хилла (кроме “Аналогии”), о чем нами было сказано ранее [10]. Но затем, в постадийном перечне доказательных методологий, на стадии II по использованию руководящих принципов, перечислены только 7 критериев, при отсутствии не только “Аналогии”, но и “Специфичности”.

22. Ключевое событие (Key Event): наблюдаемый воспроизводимый этап (который может быть измерен прямо или косвенно, или обоснованно выведен на основе знания патогенеза неблагоприятного эффекта), являющийся необходимым элементом МОА и имеющий решающее значение для результата. МОА – “Способ действия” (“Mode of Action” перевод наш. – А.К.) – биологически правдоподобная серия химически-специфичных ключевых событий, начинающаяся с воздействия и продолжающаяся через взаимодействие агента внутри клетки, а также последующие физиологические изменения и изменения тканей или органов, приводящие к неблагоприятному эффекту или исходу [108]. См. также в [10].

23. Нами была сделана попытка связаться по e-mail с первым автором [55], G. Swaen, чтобы узнать судьбу столь важной разработки, которая никак не отражалась в последующих публикациях данного исследова-

теля. Адрес был официальный, с соответствующего сайта учреждения, но ответа получено не было.

24. “Задача определения причинно-следственной связи в общественном здравоохранении всегда определялась ограниченностью доступных данных, пониманием лежащих в основе [всего] биологических или социологических процессов, и нашей способностью вмешиваться в реальный мир. Столкнувшись с иногда ограниченными данными и зачастую с плохим пониманием сети взаимосвязанных факторов в сложном мире, мы возвращаемся к прагматизму”. (“The challenge of determining causation in public health has always been shaped by the limitations of the available data, the understanding of the underlying biological or sociological processes, and our ability to intervene in the real world. Faced with sometimes limited data and an often poor understanding of a network of connected factors in a complex world, we revert to pragmatism”.) [22].

25. “If we ask for proof in medicine, or any other empirical science, we may be asking for something that does not exist” [122] (первоисточник недоступен; цитировано по [123, 124]).

26. “Контрафактический идеал” обосновывает идеальную контрольную группу, когда индивидум или группа людей, подвергавшихся воздействию, сравниваются с тем же самым индивидумом или с той же самой группой, но без воздействия. Причем – *в то же самое время*. Подобный подход на деле невозможен, поэтому используется замена на реальные группы сравнения, то есть контроли (“контрасты”), имитирующие “идеал” [7, 8].

27. “Вся научная работа является неполной, будь то наблюдательные или экспериментальные исследования. Вся научная работа может быть разрушена или изменена путем продвижения знания. [Но] это не дает нам свободы игнорировать уже имеющиеся данные или откладывать действие, которое, по-видимому, требуется в данный момент”. (“All scientific work is incomplete – whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time”.) [11].

28. “...now very few philosophers or scientists still think that scientific knowledge is, or can be, proven knowledge. But few realize that with this the whole classical structure of intellectual values falls in ruins and has to be replaced: one cannot simply water down the ideal of proven truth – as some logical empiricists do – to the ideal of “probable truth” or – as some sociologists of knowledge do – to “truth by [changing] consensus”” [132].

29. В работе Cox L., 2018 [26] причинные подходы в эпидемиологии расширены до девяти, но вероятностная причинность названа первой. Итак: 1) Вероятностная причинность; 2) Ассоциативная причинность (более высокие риски – при более высоких воздействиях); 3) Атрибутивная причинность (R дополнительных случаев в человеко-лет на единицу воздействия); 4) Контрафактическая причинность (снижение воздействия на единицу показателя обуславливает уменьшение частоты на R человеко-лет); 5) Прогнозирующая (predictive) причинность (наблюдение за эффектом после контрафактического вмешательства); 6) Структурная причинность (среднее количество случа-

ев на человеко-лет выводится из величины воздействия); 7) Манипулятивная причинность (уменьшение воздействия на одну единицу снижает количество ожидаемых случаев на R человеко-лет); 8) Объяснительная/механистическая причинность (увеличение воздействия на одну единицу вызывает изменения, распространяющиеся через биологическую сеть причинных механизмов); 9) Без причинной связи.

30. “However, although we can engage in fantasies of deterministic causal attribution, we should recognize that by ignoring the constraints imposed by how the material world is, we encourage the mythopoetics of personalized medicine. Epidemiologists, surely, should be suspicious of such” [142].

31. Явные отрицания критериев причинности в работах К.Д. Rothman вплоть до 2005 г. [52, 110, 137, 138] сглажены в последних пособиях 2008–2014 гг. [13, 53, 143]. Продолжая стоять на позиции детерминированной причинности и только гипотетико-дедуктивного метода, К.Д. Rothman (один или с соавторами) минимум с 2005 г. [52, 110] указывали, что, хотя по философским соображениям набора причинных критериев не существует, тем не менее, “списки причинных критериев стали популярными, возможно, потому, что они, как кажется, обеспечивают дорожную карту [для прохождения] через сложную территорию” [13, 52, 110, 143].

32. “The main causal model used by epidemiologists today is Rothman’s “pies”” [19].

33. “...the current system of epidemiology is very much determined by the thinking of Rothman” [86].

34. “Rothman’s 1986 book *Modern Epidemiology* [127] represented a watershed moment in the discipline” [142].

35. Как сказано в [156], “по иронии судьбы, NRC USA продолжает рекомендовать систему установления причинности GRADE, которая тесно связана с критериями Хилла [155]”.

36. “Epidemiologists seem to confuse the practical results of epidemiological research at the population level with the metaphysical views about the reality of disease causation at the individual level in their writings on causation” [158].

37. “The scientific method does not operate in isolation, but is conducted by the scientific community, which has specific internal norms to guide the activities of scientists in applying the scientific method. These norms include truthfulness, consistency, coherence, testability, reproducibility, validity, reliability, openness, impartiality and transparency” [161, 162]. Эти положения уже цитировались нами ранее [1].

38. В качестве примера условной, но научно обоснованной оценки в НКДАР-2012 [161] приводится пример с трансгенерационными эффектами облучения у человека (т.е. нарушениями и патологиями у необлученных потомков облученных родителей). Эпидемиология не дала доказательств наличия таковых за более чем 70 лет масштабных исследований (с 1948 г. [163]). Однако НКДАР сделал вывод о рисках трансгенерационных эффектов радиации для человека на основе экспериментов на животных, для которых таковые эффекты выявляются однозначно (см. в [6, 9, 118]).

39. “Incorporating non-scientific concerns. This approach may or may not take account of science-based inferences. In this case, decision-makers may take account of

norms external to science such as social responsibility, ethics, utility, prudence, precaution and practicality of application. Such considerations, while important, are outside of the Committee’s remit, and are mentioned here only for the sake of completeness” [161].

40. Однозначное акцентирование в [43, 91] на различиях выводов в исследованиях [170, 171] по влиянию алкоголя на частоту рака молочной железы не представляется столь однозначным. Использование в [170] (1989) критериев “Сила связи”, “Постоянство ассоциации”, “Специфичность”, “Временная зависимость” и “Согласованность с текущими фактами и теоретическими знаниями (Coherence)” привело к выводу об отсутствии причинности. Заключение второго исследования [171] (1990), основанное на тех же критериях, было хотя и несколько иным, но не столь уж категоричным: “при факторах риска рака молочной железы, но не сердечных заболеваний, разумно ограничить употребление алкоголя до тех пор, пока не будут изучены его эффекты”. Ныне МАИР классифицирует алкоголь как фактор риска рака молочной железы [172].

41. “Burch is very critical of the statement that the “causal significance of an association is a matter of judgment which goes beyond any statement of statistical probability”. This is true in science in general. One can place a statement of probability on the findings of a specific experiment, or of a sample survey, but it is impossible to place a probability statement and confidence limits on the truth of a scientific hypothesis. A major reason for this is that a demonstration of causation depends upon a synthesis of sets of data of different types. It would be worthwhile summarizing this logical approach, since it serves as a general background for the more detailed discussion” [129].

42. Метод проверки наличия конфаундера для ассоциации эффекта с интересующим фактором. Подыскивается иное воздействие или эффект, которое может быть обусловлено тем же конфаундером, но ассоциация с которым лишена биологического правдоподобия. Если и здесь связь сохраняется, значит, налицо присутствие единого вмешивающегося фактора. Так, курение дает одинаковую ассоциацию со смертностью и от суицидов (биологический механизм можно придумать), и от убийств (правдоподобие отсутствует). В обоих случаях вероятно влияние поведенческих факторов [178]. Другие примеры см. в [8].

43. “Scientists these days tend to keep up a polite fiction that all science is equal. Except for the work of the misguided opponent whose arguments we happen to be refuting at the time, we speak as though every scientist’s field and methods of study are as good as every other scientists, and perhaps a little better. This keeps us all cordial when it comes to recommending each other for government grants” [179].

44. Alvan R. Feinstein (США; 1925–2001): “Мета-анализ — статистическая алхимия 21 века... идея получить что-то даром, одновременно игнорируя установленные научные принципы” [181]. Был таким критиком, что даже по сути некролог ему, написанный А. Morabia в 2002 г., имеет название: “Спорная дискуссия о страстном полемисте” [182]. А.Р. Feinstein и умер, как и жил: от сердечного приступа во время дискуссии на одном из симпозиумов [182]. Данный автор ввел в эпидемиологию и доказательную медицину:

термин “золотой стандарт” для RCT (1982 г., вместе с R.I. Horwitz [183]), наименование “классическая эпидемиология” [184, 185] и др. [9].

45. “In a controlled trial, as in all experimental work, there is no need in the search for precision to throw common sense out of the window” [116].

46. Можно сделать, однако, иные подборки на любой вкус: подтверждающие “стандартное” положение прежних времен, что обсервационные исследования завывают эффекты сравнительно с RCT или, реже, что занижают их (источники см. в систематических обзорах [187–189]).

47. “These criteria offered by Hill are saddled with reservations and exceptions” (1982) [191] (цитировано по [121]) и “As is evident, the standards of epidemiologic evidence offered by Hill are saddled with reservations and exceptions” (2005–2008) [13, 52, 110, 138].

48. “Although checklists may not be appropriate for causal inference, the points laid out by Hill are still important considerations. The criteria may be useful when applied in the context of specific hypotheses” [53].

49. “Assertions that each of Hill’s criteria can be dispensed with (usually advanced authoritatively without supporting evidence) are improvident if the goal is to establish causation” [34].

50. INUS – “Insufficient and Non-redundant part of an Unnecessary but Sufficient condition” [210, 211], т.е. “Недостаточная, но не лишняя часть не необходимого, но достаточного условия”. Перевод наш, поскольку на русском языке не встретилось ничего лучше. Например, в философской статье из МГУ Д.А. Голосного от 2014 г. [212], где рассматриваются воззрения J.L. Mackie, взяты какие-то юридические обороты (“причинение”). В [212] использована конструкция: “выступает недостаточной, но необходимой частью условия, которое само по себе является не-необходимым, но достаточным”. Это не близко к оригиналу. В 1965 г. [210] J.L. Mackie представил в качестве примера INUS комплексную причинность пожара: произошло короткое замыкание, которое не является необходимым условием, поскольку к пожару могли привести и иные причины. Равным образом оно и недостаточно, так как, если бы рядом не было горючих материалов, то замыкание бы ничего не дало. Причина носила, таким образом, комплексный характер и состояла из сочетания разных факторов, которые сделали ее достаточной. В 1974 г. [211] J.L. Mackie привел пример с приемом яда, когда человек не использовал антидот и не имел природного очищения путем рвоты. Какова причина смерти? Умер ли человек от того, что в организм поступил яд, или от того, что не принял антидот, или же от того, что его желудок не сработал рвотой? Одно только поступление яда недостаточно: многие принимали яд без наступления смерти (поскольку их вырвало). Но прием яда является частью совокупности условий, которые совместно достаточны для летального исхода. Кроме того, исходя именно из этого комплекса летальных условий (не принимая во внимание другие комплексы, достаточные для смерти от иных причин), один только прием яда не является лишним: смертные случаи *от яда* не происходят при обстоятельствах, когда яд не был принят. Таким образом, прием яда – недостаточная, но и не лишняя часть не необходимого,

но достаточного условия для смерти (цитировано по [213]).

51. “In our ignorance, the best we can do in assessing risk is to classify people according to measured risk indicators and then assign the average risk observed within a class to persons within the class. As knowledge or specification of additional risk indicators expands, the risk estimates assigned to people will depart from average according to the presence or absence of other factors that predict the outcome” [13]. Это третье издание “Modern Epidemiology”, от 2008 г. Судя по цитате в [142], практически аналогичные рассуждения имели место и в первом издании от 1986 г. [137].

52. “What is required is much more than the application of a list of criteria. Instead, one must apply thorough criticism, with the goal of obtaining a quantified evaluation of the total error that afflicts the study. This type of assessment is not one that can be done easily by someone who lacks the skills and training of a scientist familiar with the subject matter and the scientific methods that were employed. Neither can it be applied readily by judges in court, nor by scientists who either lack the requisite knowledge or who do not take the time to penetrate the work” [110].

53. Согласно “Health belief model” (“Модель ожидания здоровья”; перевод наш, поскольку устоявшаяся русскоязычная терминология отсутствует), люди, как правило, не желают менять свое поведение, если они не верят в преимущества подобных изменений [228]. К этому полезно добавить цитату из Hill A.V., 1965 [11]: “чтобы заставить людей сжигать в своих домах то топливо, которое им не нравится, или заставить перестать курить, а также есть жиры и сахар, которые им нравятся”, нужны были “очень веские доказательства” (оригинал цитаты см. в [10]).

54. Jeremiah Noah Morris (1910–2009; Шотландия) – известный эпидемиолог в области здравоохранения, который впервые связал важность физической активности для предотвращения сердечно-сосудистых заболеваний. Популяционный подход; индивидуальные и групповые риски [234].

55. Для П.С. Лапласа вероятность относительна: “Вероятность обуславливается отчасти этим незнанием, а отчасти нашим знанием” (“Probability is relative, in part to ignorance, in part to our knowledge”) [238]; цитировано по [226]). Если бы мы могли знать мир во всей его полноте и располагать бесконечным интеллектом, то нам якобы не понадобилась бы вероятность [226].

56. “The conceptual framework for causes presented here is intended neither as a review nor an expansion of knowledge, but rather as a viewpoint which bridges the gap between metaphysical notions of cause and basic epidemiologic parameters. The focus, then, is neither metaphysics nor epidemiology, but the gulf between them... may facilitate communication about causes of illness” [215].

57. “...epidemiologists seem to confuse the practical results of epidemiological research at the population level with the metaphysical views about the reality of disease causation at the individual level in their writings on causation”; “On the one hand, I answer the question for a useful notion of cause for epidemiological practice by defending a probabilistic account in terms of average effects as an appropriate one. On the other hand, I answer the question of the metaphysical presuppositions of causal concepts

in epidemiology by defending Kenneth Rothman's view on causation as one providing an insightful answer to that question" [158].

58. "SCC models (the only ones originating in epidemiology) stand apart in requiring specification of mechanisms within the individual units under study. There are rarely data to support such detailed specification, which may explain why SCC models have seen little use beyond teaching examples" [12].

59. "An example is: "If sepsis had occurred, the patient would have died" and "if no sepsis had occurred, the patient would not have died". We need not go into the details of this widely acclaimed theory because its very basis is too speculative and beyond any accessible reality" [202].

60. "It is absolutely impossible to ground them in the experiential world and test them empirically. Accordingly, an advocate of a counterfactual will fail to reasonably answer the basic epistemological question "how do you know that?". Notwithstanding their intriguing metaphysical aura, counterfactuals cannot be relied upon and fruitfully used in empirical-practical domains such as medicine. Diagnostic-therapeutic decision-making ought not to rest on untestable etiologic speculation if more reliable knowledge is available. In addition, it has been convincingly demonstrated that counterfactuals lead to inferential absurdities in deductions and should therefore be avoided in scientific reasoning" [202].

61. Неспособность выявить избыточный риск смертности от рака для большинства работников мировой ядерной индустрии, накопивших за период занятости, согласно нашему исследованию (63 когорты из 18 стран и интернациональные группы), средние дозы, редко превышающие 100 мЗв [324], приводят к выводу о трудностях фиксации для подобных доз стохастических эффектов вследствие неопределенностей и конфаундеров [219].

62. Весомые ссылки с конкретными зависимостями "Возраст – частота смертности от рака" нелегко найти, поэтому мы привели все нам известные.

63. Относительно вопроса, действительно ли растет ныне частота смертности от рака от года к году (скорее, наоборот, по данным РОССТАТа), см. в нашей работе [10] примечание 32.

64. "We use a DAG to visualise complex associations, but when we only have observational data at our disposal, we must find other ways to assess a) whether a particular association is causal and not due to confounding or other bias, and b) what the direction of this association is". "One way to answer this question is by incorporating prior knowledge in Hill's scheme (or similar criteria) with DAGs to determine which edges can be considered causal" [336].

КОНФЛИКТ ИНТЕРЕСОВ И ВОЗМОЖНОСТЬ СУБЪЕКТИВНЫХ УКЛОНОВ

Конфликт интересов отсутствует. Представленное исследование, выполненное попутно в рамках более широкой бюджетной темы НИР ФМБА России, не поддерживалось никакими иными источниками финансирования. При выполнении работы не имелось временных рамок, официальных требований, ограничений, или же иных внешних объективных либо субъективных вмешивающихся факторов.

СПИСОК ЛИТЕРАТУРЫ

1. *Котеров А.Н.* Критерии причинности в медико-биологических дисциплинах: история, сущность и радиационный аспект. Сообщение 1. Постановка проблемы, понятие о причинах и причинности, ложные ассоциации // Радиационная биология. Радиоэкология. 2019. Т. 59. № 1. С. 1–32. [*Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 1. Problem statement, conception of causes and causation, false associations // Radiat. Biol. Radioecol. ("Radiation biology. Radioecology", Moscow). 2019. V. 59. № 1. P. 1–32. (In Russ. Engl. abstr.)]. *Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 1. Problem statement, conception of causes and causation, false associations // Biol. Bull. (Moscow). 2019. V. 46. № 11. P. 1458–1488. <https://doi.org/10.1134/S1062359019110165>. <https://doi.org/10.1134/S0869803119010065>
2. *Котеров А.Н.* Критерии причинности в медико-биологических дисциплинах: история, сущность и радиационный аспект. Сообщение 2. Постулаты Генле-Коха и критерии причинности неинфекционных патологий до Хилла // Радиационная биология. Радиоэкология. 2019. Т. 59. № 4. С. 341–375. [*Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 2. Henle-Koch postulates and criteria for causality of non-communicable pathologies before Hill // Radiat. Biol. Radioecol. ("Radiation biology. Radioecology", Moscow). 2019. V. 59. № 4. P. 341–375. (In Russ. Engl. abstr.)] <https://doi.org/10.1134/S0869803119040052>
3. *Котеров А.Н., Ушенкова Л.Н., Зубенкова Э.С. и др.* Сила связи. Сообщение 1. Градации относительного риска // Мед. радиология и радиац. безопасность. 2019. Т. 64. № 4. С. 5–17. [*Kotero A.N., Ushenkova L.N., Zubenkova E.S. et al.* Strength of association. Report 1. Graduation of relative risk // Medits. Radiologii Radiat. Bezopasnost ("Medical Radiology and Radiation Safety"; Moscow). 2019. V. 64. № 4. P. 5–17. (In Russ. Engl. abstr.)] https://doi.org/10.12737/article_5d1adb25725023.14868717
4. *Котеров А.Н., Ушенкова Л.Н., Молодцова Д.В. и др.* Сила связи. Сообщение 2. Градации величины корреляции // Мед. радиология и радиац. безопасность. 2019. Т. 64. № 6. С. 12–24. [*Kotero A.N., Ushenkova L.N., Molodtsova D.V. et al.* Strength of association. Report 2. Graduation of correlation size // Medits. Radiologii Radiat. Bezopasnost ("Medical Radiology and Radiation Safety"; Moscow). 2019. V. 64. № 6. P. 12–24. (In Russ. Engl. abstr.)] <https://doi.org/10.12737/1024-6177-2019-64-6-12-24>
5. *Котеров А.Н., Ушенкова Л.Н., Бирюков А.П.* Критерий Хилла "Временная зависимость". Обратная причинность и ее радиационный аспект // Радиационная биология. Радиоэкология. 2020. Т. 60. № 2. С. 115–152. [*Kotero A.N., Ushenkova L.N., Biryukov A.P.* Hill's criteria "Temporality". Reverse causation and its radiation aspect // Radiat. Biol. Radioecol. ("Radiation biology. Radioecology", Moscow). 2020. V. 60. № 2. P. 115–152. (In Russ. Engl. abstr.)]. *Kotero A.N.*,

- Ushenkova L.N., Biryukov A.P.* Hill's Temporality criterion: reverse causation and its radiation aspect // *Biol. Bull. (Moscow)*. 2020. Т. 47. № 12. С. 1–33. <https://doi.org/10.1134/S1062359020120031>. <https://doi.org/10.31857/S086980312002006X>
6. *Котеров А.Н., Ушенкова Л.Н., Бирюков А.П.* Критерий Хилла “Биологическое правдоподобие”. Интеграция данных из различных дисциплин в эпидемиологии и радиационной эпидемиологии // *Радиационная биология. Радиоэкология*. 2020. Т. 60. № 5. С. 453–480. [*Kotero A.N., Ushenkova L.N., Biryukov A.P.* Hill's criteria “Biological plausibility”. The data integration from different disciplines in Epidemiology and Radiation Epidemiology // *Radiats. Biol. Radioecol. (“Radiation biology. Radioecology”, Moscow)*. 2020. V. 60. № 5. P. 453–480. (In Russ. Engl. abstr.)] <https://doi.org/10.31857/S0869803120050069>
 7. *Котеров А.Н., Ушенкова Л.Н., Бирюков А.П.* Критерий Хилла “Эксперимент”. Контрафактический подход в дисциплинах нерадиационного и радиационного профиля // *Радиационная биология. Радиоэкология*. 2020. Т. 60. № 6. С. 565–594. [*Kotero A.N., Ushenkova L.N., Biryukov A.P.* Hill's criteria “Experiment”. The counterfactual approach in non-radiation and radiation sciences // “*Radiats. Biol. Radioecol. (“Radiation biology. Radioecology”, Moscow)*”. 2020. V. 60. № 6. P. 565–594. (In Russ. Engl. Abstr.)]. <https://doi.org/10.31857/S0869803120060193>
 8. *Котеров А.Н.* Критерии причинности в медико-биологических дисциплинах: история, сущность и радиационный аспект. Сообщение 3. Часть 1: Первые пять критериев Хилла: Использование и ограничения // *Радиационная биология. Радиоэкология*. 2021. Т. 61. № 3. С. 300–332. [*Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 3, Part 1: first five Hill's criteria: use and limitations // *Radiats. Biol. Radioecol. (“Radiation biology. Radioecology”, Moscow)*. 2021. V. 61. № 3. P. 300–332. (In Russ. Engl. abstr.)]. <https://doi.org/10.31857/S0869803121030085>
 9. *Котеров А.Н.* Критерии причинности в медико-биологических дисциплинах: история, сущность и радиационный аспект. Сообщение 3. Часть 2: последние четыре критерия Хилла: использование и ограничения // *Радиационная биология. Радиоэкология*. 2021. Т. 61. [*Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 3, Part 2: last four Hill's criteria: use and limitations // *Radiats. Biol. Radioecol. (“Radiation biology. Radioecology”, Moscow)*. 2021. V. 61. (In Russ. Engl. abstr.)]
 10. *Котеров А.Н.* Критерии причинности в медико-биологических дисциплинах: история, сущность и радиационный аспект. Сообщение 4. Часть 1: критерии после Хилла: и экологические критерии // *Радиационная биология. Радиоэкология*. 2021. Т. 61. [*Kotero A.N.* Causal criteria in medical and biological disciplines: history, essence and radiation aspect. Report 4, Part 1: Post Hill criteria and ecological Criteria // *Radiats. Biol. Radioecol. (“Radiation biology. Radioecology”, Moscow)*. 2021. V. 61. (In Russ. Engl. abstr.)]
 11. *Hill A.B.* The environment and disease: association or causation? // *Proc. R. Soc. Med.* 1965. V. 58. № 5. P. 295–300. <https://doi.org/10.1177/0141076814562718>
 12. *Greenland S., Brumback B.* An overview of relations among causal modelling methods // *Int. J. Epidemiol.* 2002. V. 31. № 5. P. 1030–1037. <https://doi.org/10.1093/ije/31.5.1030>
 13. *Rothman K.J., Greenland S., Poole C., Lash T.L.* Causation and causal inference // In: “*Modern Epidemiology*” / Eds K.J. Rothman, S. Greenland, T.L. Lash. 3rd ed. Philadelphia (PA): Wolters Kluwer, 2008. P. 5–31.
 14. *Greenland S., Pearl J.* Causal Diagrams // In: “*International Encyclopedia of Statistical Science*” / Ed. M. Lovric. 2011. P. 208–216. https://doi.org/10.1007/978-3-642-04898-2_9
 15. *Egilman D., Kim J., Biklen M.* Proving causation: the use and abuse of medical and scientific evidence inside the courtroom – an epidemiologist's critique of the judicial interpretation of the Daubert ruling // *Food Drug Law J.* 2003. V. 58. № 2. P. 223–250.
 16. *Phillips C.V., LaPole L.M.* Quantifying errors without random sampling // *BMC Med. Res. Methodol.* 2003. V. 3. Art. 9. 10 p. <https://doi.org/10.1186/1471-2288-3-9>
 17. *Phillips C.V.* Quantifying and reporting uncertainty from systematic errors // *Epidemiol.* 2003. V. 14. № 4. P. 459–466. <https://doi.org/10.1097/01.ede.0000072106.65262.ae>
 18. *Goodman K.J., Phillips C.V.* Hill's Criteria of Causation // *Encyclopedia of Statistics in Behavioral Science*. V. 1 / Eds. B.S. Everit, D.C. Howell. Chichester: John Wiley & Sons, Ltd., 2005. P. 818–820.
 19. *Vineis P., Kriebel D.* Causal models in epidemiology: past inheritance and genetic future // *Environ. Health.* 2006. V. 5. Art. 21. 10 p. <https://doi.org/10.1186/1476-069X-5-21>
 20. *Newman M.C., Zhao Y., Carriger J.F.* Coastal and estuarine ecological risk assessment: the need for a more formal approach to stressor identification // *Hydrobiologia*. 2007. V. 577. 31–40. https://doi.org/10.1007/978-1-4020-6008-3_4
 21. *Kleinberg S., Hripcsak G.* A review of causal inference for biomedical informatics // *J. Biomed. Inform.* 2011. V. 44. № 6. P. 1102–1112. <https://doi.org/10.1016/j.jbi.2011.07.001>
 22. *Glass T.A., Goodman S.N., Hernan M.A., Samet J.M.* Causal inference in public health // *Annu. Rev. Publ. Health.* 2013. V. 34. P. 61–75. <https://doi.org/10.1146/annurev-publhealth-031811-124606>
 23. *Ecological Causal Assessment* / Eds S.B. Norton, S.M. Cormier, W. Glenn. Suter II. U.S. Environmental Protection Agency. Cincinnati, OH, USA: CRC Press, 2015. 497 p.
 24. *Alexander L.K., Lopes B., Ricchetti-Masterson K., Yeatts K.B.* Causality // *Epidemiologic Research and Information Center (ERIC) Notebook*. 2nd ed. UNC Gillings School of Global Public Health, 2015. 5 p. https://sph.unc.edu/files/2015/07/nciph_ER-IC15.pdf (address data 15.01.2021)

25. *Thompson R.P., Upshur R.E.G.* Philosophy of Medicine. An Introduction. London and New York: Routledge, Taylor & Francis Group, 2018. 194 p.
26. *Cox L.* Modernizing the Bradford Hill criteria for assessing causal relationships in observational data // *Crit. Rev. Toxicol.* 2018. V. 48. № 8. P. 682–712. <https://doi.org/10.1080/10408444.2018.1518404>
27. *Szklo M., Nieto F.J.* Epidemiology. Beyond the Basics. 4th ed. Burlington: Jones & Bartlett Learning, 2019. 577 p.
28. *Martin P., Bladier C., Meek B. et al.* Weight of Evidence for hazard identification: a critical review of the literature // *Environ. Health Perspect.* 2018. V. 126. № 7. Art. 076001. <https://doi.org/10.1289/EHP3067>
29. *Pazek K., Rozman C.* Decision making under conditions of uncertainty in agriculture: a case study of oil crops // *Poljoprivreda.* 2009. V. 15. № 1. 9 p. <https://hrcak.srce.hr/39437> (address data 15.01.2021).
30. *Lilienfeld A.M.* On the methodology of investigations of etiologic factors in chronic diseases: some comments // *J. Chronic Dis.* 1959. V. 10. № 1. P. 41–46. [https://doi.org/10.1016/0021-9681\(59\)90016-5](https://doi.org/10.1016/0021-9681(59)90016-5)
31. United States Department of Health, Education and Welfare (USDHEW). Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service Publication No. 1103. Washington DC: U.S. Department of Health, Education and Welfare. 1964. 387 p. <https://biotech.law.lsu.edu/cases/tobacco/nbbmq.pdf> (address data 14.12.2020).
32. *Stallones R.A.* The association between tobacco smoking and coronary heart disease. Draft Report of June 28 to the Surgeon General's Advisory Committee on Smoking and Health. University of Minnesota Archives, Leonard M. Schuman Papers, Box 52, "Cardiovascular". 1963. (Published in: *Int. J. Epidemiol.* 2015. V. 44. № 3. P. 735–743.) <https://doi.org/10.1093/ije/dyv124>
33. *Phillips C.V., Goodman K.J.* The missed lessons of Sir Austin Bradford Hill // *Epidemiologic Perspectives & Innovations.* 2004. V. 1. Art. 3. 5 p. <https://doi.org/10.1186/1742-5573-1-3>
34. *Guzelian P.S., Victoroff M.S., Halmes N.C. et al.* Evidence-based toxicology: a comprehensive framework for causation // *Hum. Exp. Toxicol.* 2005. V. 24. № 4. P. 161–201. <https://doi.org/10.1191/0960327105ht517oa>
35. *Kundi M.* Causality and the interpretation of epidemiologic evidence // *Environ. Health Perspect.* 2006. V. 114. № 7. P. 969–974. <https://doi.org/10.1289/ehp.8297>
36. *Andersen H.* History and philosophy of modern epidemiology. Based on a talk delivered at the &HPS Conference, Pittsburgh, October 2007. <http://philsci-archive.pitt.edu/4159/> (дата обращения 15.01.2021).
37. *Howick J., Glasziou P., Aronson J.K.* The evolution of evidence hierarchies: what can Bradford Hill's "guidelines for causation" contribute? // *J. R. Soc. Med.* 2009. V. 102. № 5. P. 186–194. <https://doi.org/10.1258/jrsm.2009.090020>
38. *Biesalski H.K., Aggett P.J., Anton R. et al.* 26th Hohenheim Consensus Conference, September 11, 2010 Scientific substantiation of health claims: evidence-based nutrition // *Nutrition.* 2011. V. 27. № 10. Suppl. S1–S20. <https://doi.org/10.1016/j.nut.2011.04.002>
39. *Gordis L.* Epidemiology. 5th ed. Philadelphia: Saunders, Elsevier Inc., 2014. 392 p.
40. *Aschengrau A., Seage G.R., III.* Epidemiology in Public Health. 3rd ed. Burlington: Jones & Bartlett Learning, LLC, 2014. 596 p.; 4th ed. Burlington: Jones & Bartlett Learning, LLC, 2020. 528 p.
41. *Susser M.* Rules of inference in epidemiology // *Regul. Toxicol. Pharmacol.* 1986a. V. 6. № 2. P. 116–128. [https://doi.org/10.1016/0273-2300\(86\)90029-2](https://doi.org/10.1016/0273-2300(86)90029-2)
42. *Susser M.* The logic of Sir Karl Popper and the practice of epidemiology // *Am. J. Epidemiol.* 1986b. V. 124. № 5. P. 711–718. <https://doi.org/10.1093/oxfordjournals.aje.a114446>
43. *Weed D.L., Gorelic L.S.* The practice of causal inference in cancer epidemiology // *Cancer Epidemiol. Biomarkers Prev.* 1996. V. 5. № 4. P. 303–311.
44. *Власов В.В.* Эпидемиология: учебное пособие. 2-е изд., испр. М.: ГЭОТАР-Медиа, 2006. 464 с. [*Vlasov V.V.* Epidemiology. 2nd ed., revised. Moscow: GEOTAR-Media, 2006. 464 p. (In Russ.)]
45. USEPA 2005. Guidelines for carcinogen risk assessment. EPA/630/P-03/001B. Washington, DC: Risk Assessment Forum. National Center for Environmental Assessment Office of Research and Development U.S. Environmental Protection Agency, 2005. 166 p.
46. *Lower G.M., Kanarek M.S.* Conceptual/operational criteria of causality: relevance to systematic epidemiologic theory // *Med. Hypotheses.* 1983. V. 11. P. 217–244. [https://doi.org/10.1016/0306-9877\(83\)90064-6](https://doi.org/10.1016/0306-9877(83)90064-6)
47. *Schlesselman J.J.* "Proof" of cause and effect in epidemiologic studies: criteria for judgment // *Prev. Med.* 1987. V. 16. № 2. P. 195–210. [https://doi.org/10.1016/0091-7435\(87\)90083-1](https://doi.org/10.1016/0091-7435(87)90083-1)
48. *Charlton B.G.* Attribution of causation in epidemiology: chain or mosaic? // *J. Clin. Epidemiol.* 1996. V. 49. № 1. P. 105–107. [https://doi.org/10.1016/0895-4356\(95\)00030-5](https://doi.org/10.1016/0895-4356(95)00030-5)
49. *Kaufman J.S., Poole C.* Looking back on "causal thinking in the health sciences" // *Annu. Rev. Public Health.* 2000. V. 21. P. 101–119. <https://doi.org/10.1146/annurev.publhealth.21.1.101>
50. *Weed D.L.* Precaution, prevention, and public health ethics // *J. Medicine and Philosophy.* 2004. V. 29. № 3. P. 313–332. <https://doi.org/10.1080/03605310490500527>
51. *Hofler M.* The Bradford Hill considerations on causality: a counterfactual perspective // *Emerg. Themes Epidemiol.* 2005. V. 2. № 11. 9 p. <https://doi.org/10.1186/1742-7622-2-11>
52. *Rothman K., Greenland S.* Hill's Criteria for Causality // *Encyclopedia of Biostatistics, Online.* John Wiley & Sons, Ltd., 2005. 4 p. <https://www.rtihs.org/sites/default/files/26902%20Rothman%201998%20The%20encyclopedia%20of%20biostatistics.pdf> (дата обращения 11.10.2020).
53. *Rothman K.J.* Epidemiology. An Introduction. 2nd ed. Oxford University Press Inc., 2012. 268 p.

54. *Phillips C.V., Goodman K.J.* Hill's considerations for causal inference // *Encyclopedia of Epidemiology*. Two Volume Set / Ed. S. Boslaugh. Saint Louis University, SAGE Publications, Inc., 2008. P. 494–495.
55. *Swaen G., van Amelsvoort L.* A Weight of Evidence approach to causal inference // *J. Clin. Epidemiol.* 2009. V. 62. № 3. P. 270–277. <https://doi.org/10.1016/j.jclinepi.2008.06.013>
56. *Webb P., Bain C.* *Essential Epidemiology. An Introduction for Students and Health Professionals*. 2nd ed. Cambridge etc.: Cambridge University Press, 2011. 445 p.
57. *Friis R.H., Sellers T.A.* *Epidemiology for Public Health Practice*. 5th ed. Burlington: Jones & Bartlett Learning, 2014. 804 p.
58. *Clinical Epidemiology. Practice and Methods*. 2nd ed. / Eds P.S. Parfrey, B.J. Barrett. New York: Humana Press (brand of Springer), 2015. 533 p.
59. *Bhopal R.S.* *Concepts of Epidemiology: Integrated the ideas, theories, principles and methods of epidemiology*. 3rd ed. Oxford: University Press, 2016. 442 p.
60. *Bruce N., Pope D., Stanistreet D.* *Quantitative Methods for Health Research. A Practical Interactive Guide to Epidemiology and Statistics*. 2nd ed. Oxford: John Wiley & Sons, 2019. 545 p.
61. USEPA 1998. *Guidelines for Carcinogen Risk Assessment*. EPA/630/R-95/002F. U.S. Environmental Protection Agency. Washington, DC: Risk Assessment Forum, 1998. 188 p.
62. *Hollingsworth J.G., Lasker E.G.* The Case against differential diagnosis: Daubert, medical causation. Testimony, and the scientific method // *J. Health Law*. 2004. V. 37. № 1. P. 85–111.
63. *Parascandola M.* Two approaches to etiology: the debate over smoking and lung cancer in the 1950s // *Endeavour*. 2004. V. 28. № 2. P. 81–86. <https://doi.org/10.1016/j.endeavour.2004.02.003>
64. *Ward A.C.* The role of causal criteria in causal inferences: Bradford Hill's "aspects of association" // *Epidemiol. Perspect. Innov.* 2009. V. 6. № 2. 22 p. <https://doi.org/10.1186/1742-5573-6-2>
65. UNSCEAR 2012. Report to the General Assembly, with Scientific Annex. Annex B. Uncertainties in risk estimates for radiation induced cancer. New York, 2014. 219 p.
66. *Fedak K.M., Bernal A., Capshaw Z.A., Gross S.* Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology // *Emerg. Themes Epidemiol.* 2015. V. 12. Article 14. <https://doi.org/10.1186/s12982-015-0037-4>
67. *Olsen J., Jensen U.J.* Causal criteria: time has come for a revision // *Eur. J. Epidemiol.* 2019. V. 34. № 6. P. 537–541. <https://doi.org/10.1007/s10654-018-00479-x>
68. *Morabia A.* History of epidemiological methods and concepts // *Handbook of Epidemiology*. 2nd ed. / Ed. W. Ahrens, I. Pigeot. New York, Heidelberg, Dordrecht, London: Springer, 2014. P. 43–74.
69. *Hill A.B. Bradford.* Principles of Medical Statistics. 6th ed. London: The Lancet limited. 1955. *Хилл Бредфорд А.* Основы медицинской статистики. 6 издание: Пер. с англ. Л.К. Хоцянова / Под ред. А.М. Меркова. М.: Медгиз, 1958.306 с.
70. *Hill A.B.* Suspended judgment. Memories of the British streptomycin trial in tuberculosis: the first randomized clinical trial // *Control. Clin. Trials*. 1990. V. 11. № 2. P. 77–79. [https://doi.org/10.1016/0197-2456\(90\)90001-i](https://doi.org/10.1016/0197-2456(90)90001-i)
71. *Vineis P.* Causality assessment in epidemiology // *Theor. Med.* 1991. V. 12. № 2. P. 171–181. <https://doi.org/10.1007/BF00489797>
72. Smoking and reproductive life. The impact of smoking on sexual, reproductive and child health / Eds D. Carter, N. Nathanson, C. Seddon et al. British Medical Association. Board of Science and Education & Tobacco Control Resource Centre, 2004. 64 p. https://www.rauchfrei-info.de/fileadmin/main/data/Dokumente/Smoking_ReproductiveLife.pdf (address data 16.01.2021).
73. The Health Consequences of Smoking: A Report of the Surgeon General. Rockville, MD: Office of the Surgeon General, US Public Health Service, 2004. 910 p.
74. *Lagiou P., Trichopoulos D., Adami H.O.* Concepts in cancer epidemiology and etiology // *Textbook of Cancer Epidemiology*. 2nd ed. / Eds H.O. Adami, D. Hunter, D. Trichopoulos. New York etc.: Oxford University Press, 2008. P. 127–152.
75. *Rhomberg L.R., Goodman J.E., Bailey L.A. et al.* A survey of frameworks for best practices in Weight-of-Evidence analyses // *Crit. Rev. Toxicol.* 2013. V. 43. № 9. P. 753–784. <https://doi.org/10.3109/10408444.2013.832727>
76. *Weed D.L.* Causal criteria and Popperian refutation // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 15–32.
77. *Cormier S.M., Suter G.W. II, Norton S.B.* Causal characteristics for ecoepidemiology // *Hum. Ecol. Risk Assess.* 2010. V. 16. № 1. P. 53–73. <https://doi.org/10.1080/10807030903459320>
78. *Verbeek J.* When work is related to disease, what establishes evidence for a causal relation? // *Saf. Health Work.* 2012. V. 3. № 2. P. 110–116. <https://doi.org/10.5491/SHAW.2012.3.2.110>
79. *Feinstein A.R.* Clinical biostatistics. XLVII. Scientific standards vs. statistical associations and biologic logic in the analysis of causation // *Clin. Pharmacol. Ther.* 1979a. V. 25. № 4. P. 481–492. <https://doi.org/10.1002/cpt1979254481>
80. *Feinstein A.R.* Efficacy of different research structures in preventing bias in the analysis of causation // *Clin. Pharmacol. Ther.* 1979b. V. 26. № 1. P. 129–141. <https://doi.org/10.1002/cpt1979261129>
81. *Goodman S.N., Samet J.M.* Cause and Cancer Epidemiology // *Schottenfeld and Fraumeni Cancer Epidemiology and Prevention*. 3rd ed. / Eds D. Schottenfeld and J.F. Fraumeni. New York: Oxford University Press. 2006. P. 3–9; 4th ed. / Eds M.J. Thun, M.S. Linet, J.R. Cerhan, C. Haiman, D. Schottenfeld. New York: Oxford University Press. Printed by Sheridan Books, Inc., USA, 2018. P. 97–104.
82. *Thygesen L.C., Andersen G.S., Andersen H.* A philosophical analysis of the Hill criteria // *J. Epidemiol. Community Health*. 2005. V. 59. № 6. P. 512–516. <https://doi.org/10.1136/jech.2004.027524>

83. *Suppes P.* A Probabilistic Theory of Causation. Amsterdam: North-Holland, 1970. 130 p.
84. *Wulff H., Pedersen S.A., Rosenberg R.* Philosophy of Medicine. An Introduction. Oxford: Blackwell Scientific Publications, 1986. 222 p.
85. *Renton A.* Epidemiology and causation: a realist view // *J. Epidemiol. Community Health.* 1994. V. 48. № 1. P. 79–85.
<https://doi.org/10.1136/jech.48.1.79>
86. *Frick U., Rehm J.* Can we establish causality with statistical analyses? The example of Epidemiology // *Statistics and Causality: Methods for Applied Empirical Research* / Eds W. Wiedermann, A. von Eye. Hoboken, New Jersey: John Wiley & Sons, Inc., 2016. P. 407–431.
87. *Russo F., Williamson J.* Interpreting causality in the health sciences // *Int. Stud. Philos. Sci.* 2007. V. 21. № 2. P. 157–170.
<https://doi.org/10.1080/02698590701498084>
88. *Howick J., Glasziou P., Aronson J.K.* Evidence-based mechanistic reasoning // *J. R. Soc. Med.* 2010. V. 103. № 11. P. 433–441. .
<https://doi.org/10.1258/jrsm.2010.100146433>
89. *Fox G.A.* Practical causal inference for ecoepidemiologists // *J. Toxicol. Environ. Health.* 1991. V. 33. № 4. P. 359–373.
<https://doi.org/10.1080/15287399109531535>
90. *Weed D.L.* Epidemiologic evidence and causal inference // *Hematol. Oncol. Clin. North Am.* 2000. V. 14. № 4. P. 797–807.
[https://doi.org/10.1016/S0889-8588\(05\)70312-9](https://doi.org/10.1016/S0889-8588(05)70312-9)
91. *Weed D.L.* On the use of causal criteria // *Int. J. Epidemiol.* 1997. V. 26. № 6. P. 1137–1141.
<https://doi.org/10.1093/ije/26.6.1137>
92. National Research Council (US). Committee on Diet and Health. Diet and health: implications for reducing chronic disease risk. Washington, DC: National Academy Press, 1989.
<https://doi.org/10.17226/1222>
93. *Coughlin S.S.* Causal Inference and Scientific Paradigms in Epidemiology. Bentham E-book. 2010. 70 p.
<https://doi.org/10.2174/97816080518161100101>
<https://ebooks.benthamscience.com/book/9781608051816/> (address data 19.12.2021).
94. *Fletcher R.H., Fletcher S.W., Wagner E.H.* Clinical Epidemiology: The Essentials. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 1996. 276 p.; *Флетчер Р., Флетчер С., Вагнер Э.* Клиническая эпидемиология. Основы доказательной медицины: Пер. с англ. М.: Медиа Сфера, 1998. 352 с.
95. *Fletcher R.H., Fletcher S.W.* Clinical Epidemiology. The essential. 4th ed. Baltimore: Lippincott Williams & Wilkins. 2005. 252 p.
96. *Weed D.L.* Causation: an epidemiologic perspective (in five parts) // *J. Law & Policy.* 2003. V. 12. № 1. P. 43–53.
<http://brooklynworks.brooklaw.edu/jlp/vol12/iss1/3> (address data 19.01.2021).
97. *Weed D.L.* Weight of Evidence: a review of concept and methods // *Risk Anal.* 2005. V. 25. № 6. P. 1545–1557.
<https://doi.org/10.1111/j.1539-6924.2005.00699.x>
98. *Bonita R., Beaglehole R., Kjellstrom T.* Basic epidemiology. 2nd ed. World Health Organization, 2006. 212 p.
99. *Gordis L., Kleinman J.C., Klerman L.V. et al.* Criteria for evaluating evidence regarding the effectiveness of prenatal interventions // *New Perspectives on Prenatal Care* / Eds I.R. Merkatz, J.E. Thompson. New York: Elsevier, 1990. P. 31–38.
100. *Jorgensen T.J.* Strange Glow. The Story of Radiation. Princeton and Oxford: Princeton University Press, 2016. 490 p.
101. *Holman C.D., Arnold-Reed D.E., de Klerk N. et al.* A psychometric experiment in causal inference to estimate evidential weights used by epidemiologists // *Epidemiology.* 2001. V. 12. № 2. P. 246–255.
<https://doi.org/10.1097/00001648-200103000-00019>
102. Department of Clinical Epidemiology and Biostatistics. McMaster University Health Sciences Centre. How to read clinical journals. IV. To determine etiology or causation // *Can. Med. Assoc. J.* 1981. V. 124. № 5. P. 985–990.
103. *Sackett D.L.* The diagnosis of causation // *Epidemiological Issues in Reported Drug-induced Illnesses – S.M.O.N. and Other Examples* / Eds M. Gent, I. Shigematsu. Hamilton, Ontario: McMaster Univ. Lib. Press, 1978. P. 106–117.
104. *Burton G.A. Jr., Chapman P.M., Smith E.P.* Weight of Evidence approaches for assessing ecosystem impairment // *Human Ecol. Risk Assess.* 2002. V. 8. № 7. P. 1657–1674.
<https://doi.org/10.1080/20028091057547>
105. *Linkov I., Loney D., Cormier S. et al.* Weight-of-Evidence evaluation in environmental assessment: review of qualitative and quantitative approaches // *Sci. Total Environ.* 2009. V. 407. № 19. P. 5199–5205.
<https://doi.org/10.1016/j.scitotenv.2009.05.004>
106. *Meek M.E., Palermo C.M., Bachman A.N. et al.* Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of Weight of Evidence // *J. Appl. Toxicol.* 2014. V. 34. № 6. P. 595–606.
<https://doi.org/10.1002/jat.2984>
107. USEPA 2000. Stressor Identification Guidance Document. EPA/822/B-00-025. U.S. Environmental Protection Agency. Office of Water and Office of Research and Development. Washington, DC, USA, 2000. 246 p.
108. *Becker R.A., Dellarco V., Seed J. et al.* Quantitative Weight of Evidence to assess confidence in potential modes of action // *Regul. Toxicol. Pharmacol.* 2017. V. 86. P. 205–220.
<https://doi.org/10.1016/j.yrtph.2017.02.017>
109. *Collier Z.A., Gust K.A., Gonzalez-Morales B. et al.* A Weight of Evidence assessment approach for adverse outcome pathways // *Regul. Toxicol. Pharmacol.* 2016. V. 75. P. 46–57.
<https://doi.org/10.1016/j.yrtph.2015.12.014>
110. *Rothman K.J., Greenland S.* Causation and causal inference in epidemiology // *Am. J. Public Health.* 2005. V. 95. Suppl. 1. P. S144–S150.
111. *Gosling J.P., Hart A., Owen H.* A Bayes Linear Approach to Weight-of-Evidence risk assessment for skin allergy // *Bayesian Analysis.* 2013. V. 8. № 1. P. 169–186.
<https://doi.org/10.1214/13-BA807>
112. *Kaptschuk T.J.* The double-blind, randomized, placebo-controlled trial: gold standard or golden calf? // *J. Clin. Epidemiol.* 2001. V. 54. № 6. P. 541–549.
[https://doi.org/10.1016/s0895-4356\(00\)00347-4](https://doi.org/10.1016/s0895-4356(00)00347-4)

113. *Grossman J., Mackenzie F.* The randomized controlled trial: gold standard, or merely standard? // *Perspect. Biol. Med.* 2005. V. 48. № 4. P. 516–534. <https://doi.org/10.1353/pbm.2005.0092>
114. IARC 1987. International Agency for Research on Cancer. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Supplement 7, Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs, Vol. 1–42. Lyon, 1987. 449 p.
115. IARC 2012. Radiation. A review of human carcinogens. Vol. 100 D. IARC monographs on the evaluation of carcinogenic risks to humans. Lyon, France, 2012. 341 p.
116. *Hill A.B.* Medical ethics and controlled trials // *Br. Med. J.* 1963. V. 1. № 5337. P. 1043–1049. <https://doi.org/10.1136/bmj.1.5337.1043>
117. *Black N.* Why we need observational studies to evaluate the effectiveness of health care // *Br. Med. J.* 1996. V. 312. № 7040. P. 1215–1218. <https://doi.org/10.1136/bmj.312.7040.1215>
118. National Research Council (NRC), Division on Earth and Life Studies, Board on Radiation Effects Research, Committee to Assess Health Risks from Exposure to Low Levels of Ionizing Radiation. Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII. Phase 2. National Academies Press, 2006. 422 p.
119. *Cleland C.E.* Methodological and epistemic differences between historical science and experimental science // *Philosophy of Science.* 2002. V. 69. № 3. P. 447–451. <https://doi.org/10.1086/342455>
120. *Greenland S.* Probability versus Popper. An elaboration of the insufficiency of current Popperian approaches for epidemiologic analysis // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 95–104.
121. *Lanes S.F.* The logic of causal inference // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 59–75.
122. *Cornfield J.* Statistical relationships and proof in medicine [Editorial] // *Am. Statistician.* 1954. V. 8. № 5. P. 19–21.
123. *Lanes S.F.* Error and uncertainty in causal inference // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 173–188.
124. *Onn L.T.* A manual on how to do research // Workshop “Research methods and medical statistics”. Kuala Lumpur Hospital. Department of Nephrology, 2016. 151 p. https://www.crc.gov.my/wp-content/uploads/2016/07/03_Research_method_and_medical_statistics.pdf (address data 02.02.2021).
125. *Hume D.* A Treatise of Human Nature. 2nd ed. Oxford: Oxford University Press, 1978. [Юм. Д. Сочинения в 2-х томах. Т. 1: Пер. с англ. С.И. Церетели и др. 2-е изд., дополн. и испр. М.: Мысль, 1996. 733 с.]
126. *Popper K.R.* The Logic of Scientific Discovery. London and New York: Routledge Classics, 2002. 513 p.
127. *Lipton R., Odegaard T.* Causal thinking and causal language in epidemiology: it’s in the details // *Epidemiol. Perspect. Innov.* 2005. V. 29. № 2. Article 8. 9 p. <https://doi.org/10.1186/1742-5573-2-8>
128. *Banegas J.R., Artalejo F.R., Calero J. del Rey.* Popper and the problem of induction in epidemiology // *Rev. Esp. Salud. Publica.* 2000. V. 74. № 4. P. 327–339. (In Spanish. English abstract.)
129. *Lilienfeld A.M.* The surgeon general’s “epidemiologic criteria for causality”: a criticism of Burch’s critique // *J. Chronic Dis.* 1983. V. 36. № 12. P. 837–845. [https://doi.org/10.1016/0021-9681\(83\)90004-8](https://doi.org/10.1016/0021-9681(83)90004-8)
130. *Blackburn H., Labarthe D.* Stories from the evolution of guidelines for causal inference in epidemiologic associations: 1953–1965 // *Am. J. Epidemiol.* 2012. V. 176. № 12. P. 1071–1077. <https://doi.org/10.1093/aje/kws374>
131. *Lilienfeld’s Foundations of Epidemiology.* 4th ed. / Original: Ed. A.M. Lilienfeld / Eds D. Schneider, D.E. Lilienfeld. New York: Oxford University Press, 2015. 333 p.
132. *Lakatos I.* Falsification and the Methodology of Scientific Research Programmes // *Can Theories be Refuted?* / Ed. S.G. Harding. Synthese Library. V. 81. Dordrecht: Springer, 1976. P. 205–259.
133. *Parascandola M., Weed D.L.* Causation in epidemiology // *J. Epidemiol. Community Health.* 2001. V. 55. № 12. P. 905–912. <https://doi.org/10.1136/jech.55.12.905>
134. *Illari P.M., Russo F., Williamson J.* Why look at causality in the sciences? A manifesto // *Causality in the Sciences* / Eds P.M. Illari, F. Russo, J. Williamson. New York: Oxford University Press, 2011. P. 11–28 (PDF page numbers; book without page numbers). <https://doi.org/10.1093/acprof:oso/9780199574131.003.0001>
135. *Kincaid H.* Causal modelling, mechanism, and probability in epidemiology // *Causality in the Sciences* / Eds P.M. Illari, F. Russo, J. Williamson. New York: Oxford University Press, 2011. P. 72–91 (PDF page numbers; book without page numbers). <https://doi.org/10.1093/acprof:oso/9780199574131.003.0004>
136. *Stehbens W.E.* The concept of cause in disease // *J. Chronic. Dis.* 1985. V. 38. P. 947–950. [https://doi.org/10.1016/0021-9681\(85\)90130-4](https://doi.org/10.1016/0021-9681(85)90130-4)
137. *Rothman K.* Modern Epidemiology. 1st ed. Boston: Little Brown MA., 1986. 358 p.
138. *Rothman K.J., Greenland S.* Causation and causal inference // *Modern Epidemiology.* 2nd ed. / Ed. K.J. Rothman. Philadelphia: Lippincott Williams & Wilkins, 1998. P. 7–28.
139. *Pearl J., Mackenzie D.* The Book of Why: The New Science of Cause and Effect. New York: Basic Books, 2018. 432 p.
140. *Gori G.B.* Epidemiologic evidence in public and legal policy: reality or metaphor? *Critical Legal Issues.* Washington: Washington Legal Foundation. Working Paper Series No. 124, 2004. 33 p.
141. *Parascandola M.* Causes, risks, and probabilities: probabilistic concepts of causation in chronic disease epidemiology // *Prev. Med.* 2011. V. 53. № 4–5. P. 232–234. <https://doi.org/10.1016/j.ypmed.2011.09.007>
142. *Davey Smith G.* Post-modern epidemiology: when methods meet matter // *Am. J. Epidemiol.* 2019. V. 188. № 8. P. 1410–1419. <https://doi.org/10.1093/aje/kwz064>

143. Rothman K.J., Greenland S. Basic concepts // Handbook of Epidemiology. 2nd ed. / Eds W. Ahrens, I. Pigeot. New York, Heidelberg, Dordrecht, London: Springer, 2014. P. 75–122.
144. Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology // Am. J. Epidemiol. 1991. V. 133. № 7. P. 635–648. <https://doi.org/10.1093/oxfordjournals.aje.a115939>
145. Susser M. Judgement and causal inference: criteria in epidemiologic studies // Am. J. Epidemiol. 1977. V. 105. № 1. P. 1–15. Reprint: Am. J. Epidemiol. 1995. V. 141. № 8. P. 701–715. <https://doi.org/10.1093/oxfordjournals.aje.a117491>
146. Susser M. Falsification, verification and causal inference in epidemiology: reconsiderations in the light of sir Karl Popper's philosophy // Causal Inference / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources. 1988. P. 33–57.
147. Weed D.L., Trock B.J. Criticism and the growth of epidemiologic knowledge. (Re: Popperian refutation in epidemiology) // Am. J. Epidemiol. 1986. V. 123. № 6. P. 1119–1121. <https://doi.org/10.1093/oxfordjournals.aje.a114341>
148. Weed D.L. Criticism and its constraints: a self-appraisal and rejoinder // Causal Inference / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 201–207.
149. Weed D.L. Truth, epidemiology, and general causation // Brook. L. Rev. 2008. V. 73. № 3. Article 4. P. 943–957. <https://brooklyn-works.brooklaw.edu/blr/vol73/iss3/4> (address data 26.12.2020).
150. Poole C. Induction does not exist in epidemiology, either // Causal Inference / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 153–162.
151. Weed D.L. An epidemiological application of Popper's method // J. Epidemiol. Community Health. 1985. V. 39. № 4. P. 277–285. <https://doi.org/10.1136/jech.39.4.277>
152. Weed D.L. On the logic of causal inference // Am. J. Epidemiol. 1986. V. 123. № 6. P. 965–979. <https://doi.org/10.1093/oxfordjournals.aje.a114349>
153. Spirtes P., Glymour C., Scheines R. Causation, Prediction, and Search. New York, NY: Springer-Verlag, 1993. 525 p.
154. Epidemiology: Principles and Practical Guidelines / Eds J. Van den Broeck, J.R. Brestoff. Dordrecht: Springer, 2013. 621 p.
155. NRC 2014. Review of EPA's integrated risk information system (IRIS) process // National Research Council. Washington, DC: The National Academies Press, 2014. 170 p.
156. Linkov I., Massey O., Keisler J. et al. From "Weight of Evidence" to quantitative data integration using multi-criteria decision analysis and Bayesian methods // ALTEX. 2015. V. 32. № 1. P. 3–8. <https://doi.org/10.14573/altex.1412231>
157. UNSCEAR 2006. Report to the General Assembly, with Scientific Annexes. Annex A. Epidemiological studies of radiation and cancer. United Nations. New York, 2008. P. 17–322.
158. De Vreese L. Epidemiology and causation // Med. Health Care Philos. 2009. V. 12. № 3. P. 345–353. <https://doi.org/10.1007/s11019-009-9184-0>
159. Causal Analysis/Diagnosis Decision Information System (CADDIS). Vol. 1. About Causal Assessment. A Conceptual and Historical Explanation of Our Causal Approach. U.S. Environmental Protection Agency, 2014. <https://www.epa.gov/caddis-vol1/about-causal-assessment> (address data 03.02.2021).
160. Rothman K.J. Inferring causal connection – habit, faith or logic? // Causal Inference / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 3–12.
161. UNSCEAR 2012. Report to the General Assembly, with Scientific Annexes. Annex A. Attributing health effects to ionizing radiation exposure and inferring risks. New York, 2015. 86 p.
162. Tranoy K.E. Science and ethics. Some of the main principles and problems // The Moral Import of Science: Essays on Normative Theory, Scientific Activity and Wittgenstein / Ed. A.J.I. Jones. Bergen: Sigma Forlag, 1988. P. 111–136.
163. Nakamura N. Genetic effects of radiation in atomic-bomb survivors and their children: past, present and future // J. Radiat. Res. 2006. V. 47. Suppl. B. P. B67–B73. <https://doi.org/10.1269/jrr.47.b67>
164. NCRP Report No 136. Evaluation of the linear-non-threshold dose-response model for ionizing radiation. National Council on Radiation Protection and Measurements, 2001. 263 p.
165. Phillips C.V., Goodman K.J. Causal criteria and counterfactuals; nothing more (or less) than scientific common sense // Emerg. Themes Epidemiol. 2006. V. 3. Article 5. 7 p. <https://doi.org/10.1186/1742-7622-3-5>
166. Rose S.P.R. Molecules and Minds. Philadelphia: Milton Keynes, Open University Press, 1987. 144 p.
167. Charlton B.G. A critique of biological psychiatry // Psychol. Med. 1990. V. 20. № 1. P. 3–6.
168. Fujimara J.H., Chou D.Y. Dissent in science: Styles of scientific practice and the controversy over the cause of AIDS // Social. Sci. Mecl. 1994. V. 38. № 8. P. 1017–1036.
169. Burch P.R. The surgeon general's "epidemiologic criteria for causality". A critique // J. Chronic Dis. 1983. V. 36. № 12. P. 821–836. [https://doi.org/10.1016/0021-9681\(83\)90003-6](https://doi.org/10.1016/0021-9681(83)90003-6)
170. Wynder E.L., Harris R.E. Does alcohol consumption influence the risk of developing breast cancer? Two views // In: "Important Advances in Oncology" / Eds V.T. DeVita, S. Hellman, and S.A. Rosenberg. Philadelphia: J.B. Lippincott Co., 1989. P. 283–293.
171. Hiatt R.A. Alcohol consumption and breast cancer // Med. Oncol. Tumor Pharmacother. 1990. V. 7. № 2–3. P. 143–151. <https://doi.org/10.1007/BF02988542>
172. Liu Y., Nguyen N., Colditz G.A. Links between alcohol consumption and breast cancer: a look at the evidence // Womens Health (Lond). 2015. V. 11. № 1. P. 65–77. <https://doi.org/10.2217/whe.14.62>
173. Rhomberg L., Bailey L.G., Hamade A., Mayfield D. Is exposure to formaldehyde in air causally associated

- with leukemia? A hypothesis-based Weight-of-Evidence analysis // *Crit. Rev. Toxicol.* 2011. V. 41. № 7. P. 555–621.
<https://doi.org/10.3109/10408444.2011.560140>
174. *Labarthe D.M., Stallones R.A.* Epidemiologic inference // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 119–129.
 175. *Young T.K.* Assessing health risk in populations // *Young T.K. Population health: concepts and methods.* New York: Oxford University Press, 1998. P. 143–49.
 176. *MacMahon B., Trichopoulos D.* Epidemiology: Principles and Methods. 2nd ed. Boston: Little, Brown and Company, 1996. 347 p.
 177. *Bhopal R.S.* Concepts of Epidemiology: an Integrated Introduction to the Ideas, Theories, Principles and Methods of Epidemiology. Oxford: University Press, 2002. 317 p.
 178. *Gage S.H., Munafo M.R., Davey Smith G.* Causal inference in developmental origins of health and disease (DOHaD) research // *Annu. Rev. Psychol.* 2016. V. 67. P. 567–585.
<https://doi.org/10.1146/annurev-psych-122414-033352>
 179. *Platt J.R.* Strong inference // *Science.* 1964. V. 146. № 3642. P. 347–353.
<https://doi.org/10.1126/science.146.3642.347>
 180. *Guyton K.Z., Barone S., Jr., Brown R.C. et al.* Mode of Action frameworks: a critical analysis // *J. Toxicol. Environ. Health B Crit. Rev.* 2008. V. 11. № 1. P. 16–31.
<https://doi.org/10.1080/10937400701600321>
 181. *Feinstein A.R.* Meta-analysis: statistical alchemy for the 21st century // *J. Clin. Epidemiol.* 1995. V. 48. № 1. P. 71–79.
[https://doi.org/10.1016/0895-4356\(94\)00110-c](https://doi.org/10.1016/0895-4356(94)00110-c)
 182. *Morabia A.* The controversial controversy of a passionate controversialist // *J. Clin. Epidemiol.* 2002. V. 55. № 12. P. 1207–1213.
[https://doi.org/10.1016/s0895-4356\(02\)00526-7](https://doi.org/10.1016/s0895-4356(02)00526-7)
 183. *Feinstein A.R., Horwitz R.I.* Double standards, scientific methods, and epidemiologic research // *N. Engl. J. Med.* 1982. V. 307. № 260. P. 1611–1617.
<https://doi.org/10.1056/NEJM198212233072604>
 184. *Feinstein A.R.* Clinical epidemiology. I. The population experiments of nature and of man in human illness // *Ann. Intern. Med.* 1968. V. 69. № 4. P. 807–820.
<https://doi.org/10.7326/0003-4819-69-4-807>
 185. *Feinstein A.R.* Clinical Epidemiology: The Architecture of Clinical Research. Philadelphia etc.: W.B. Saunders Company, 1985. 812 p.
 186. *Patlewicz G., Simon T.W., Rowlands J.C. et al.* Proposing a scientific confidence framework to help support the application of adverse outcome pathways for regulatory purposes // *Regul. Toxicol. and Pharmacol.* 2015. V. 71. № 3. P. 463–477.
<https://doi.org/10.1016/j.yrtph.2015.02.011>
 187. *Concato J., Shah N., Horwitz R.I.* Randomized, controlled trials, observational studies, and the hierarchy of research designs // *N. Engl. J. Med.* 2000. V. 342. № 25. P. 1887–1892.
<https://doi.org/10.1056/NEJM200006223422507>
 188. *Vandenbroucke J.P.* Why do the results of randomised and observational studies differ? // *Brit. Med. J.* 2011. V. 343. Article d7020. 2 p.
<https://doi.org/10.1136/bmj.d7020>
 189. *Golder S., Loke Y.K., Bland M.* Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview // *PLoS Med.* 2011. V. 8. № 5. Article e1001026.
<https://doi.org/10.1371/journal.pmed.1001026>
 190. *Hill A.B.* Reflections on the controlled trial // *Ann. Rheum. Dis.* 1966. V. 25. № 2. P. 107–113.
<https://doi.org/10.1136/ard.25.2.107>
 191. *Rothman K.J.* Causation and causal inference // *Cancer Epidemiology and Prevention* / Eds D. Schottenfeld, J.F. Fraumeni. Philadelphia: W.B. Saunders Company, 1982. P. 20–21.
 192. The Health Consequences of Smoking: A Report of the Surgeon General. Rockville, MD: Office of the Surgeon General, US Public Health Service, 1982.
 193. *Brownlee K.A.* A review of “Smoking and Health” // *J. Am. Stat. Ass.* 1965. V. 60. № 311. P. 722–739.
<https://doi.org/10.1080/01621459.1965.10480823>
 194. *Lanes S.F., Poole C.* “Truth in packaging?” The unwrapping of epidemiologic research // *J. Occup. Med.* 1984. V. 26. № 8. P. 571–574.
 195. *Ioannidis J.P.* Exposure-wide epidemiology: revisiting Bradford Hill // *Statist. Med.* 2016. V. 35. № 11. P. 1749–1762.
<https://doi.org/10.1002/sim.6825>
 196. *Cox D.R., Wermuth N.* Multivariate Dependencies. Models, Analyses and Interpretation. London: Chapman and Hall, 1996. 272 p.
 197. *Susser M.* Rational science versus a system of logic // *Causal Inference* / Ed. K.J. Rothman. Chestnut Hill, MS: Epidemiologic Resources, 1988. P. 189–199.
 198. *Karhausen L.R.* Causation: the elusive grail of epidemiology // *Med. Health Care Philos.* 2000. V. 3. № 1. P. 59–67.
<https://doi.org/10.1023/a:1009970730507>
 199. *Rockhill B.* Theorizing about causes at the individual level while estimating effects at the population level: implications for prevention // *Epidemiology.* 2005. V. 16. № 1. P. 124–129.
<https://doi.org/10.1097/01.ede.0000147111.46244.41>
 200. *McGwin G., Jr.* Causation in Epidemiology // *Am. J. Ophthalmol.* 2010. V. 150. № 5. P. 599–601.
<https://doi.org/10.1016/j.ajo.2010.06.031>
 201. International Encyclopedia of Statistical Science / Ed. M. Lovric. Berlin, New York: Springer, 2011. 1673 p. Online: 2014.
 202. *Sadegh-Zadeh K.* Handbook of Analytic Philosophy of Medicine. 2nd ed. Philosophy and Medicine. V. 119. Dordrecht etc: Springer, 2015. 1224 p.
 203. *Bahrami H.* Causal models in primary open angle glaucoma // *Ophthalmic Epidemiol.* 2006. V. 13. № 4. P. 291–298.
<https://doi.org/10.1080/09286580600803677>
 204. *Zielinski A.* Causality in epidemiological research // *Przegl. Epidemiol.* 2009. V. 63. № 4. P. 557–564.
 205. *Dumas O., Siroux V., Le Moual N., Varraso R.* Causal analysis approaches in epidemiology // *Rev. Epidemiol. Sante Publique.* 2014. V. 62. № 1. P. 53–63. (In

- French. English abstract.)
<https://doi.org/10.1016/j.respe.2013.09.002>
206. Suzuki E., Komatsu H., Yorifuji T. et al. Causal inference in medicine. Part II. Directed acyclic graphs—a useful method for confounder selection, categorization of potential biases, and hypothesis specification // *Nihon Eiseigaku Zasshi*. 2009. V. 64. № 4. P. 796–805. (In Japanese. English abstract.)
<https://doi.org/10.1265/jjh.64.796>
207. Suzuki E., Tsuda T., Yamamoto E. Covariate balance for no confounding in the sufficient-cause model // *Ann. Epidemiol.* 2018. V. 28. № 1. P. 48–53.
<https://doi.org/10.1016/j.annepidem.2017.11.005>
208. Hulswit M. A short history of “Causation” // *Cause to Causation*. Springer Science & Business Media, 2002. 258 p. <http://see.library.utoronto.ca/SEED/Vol4-3/Hulswit.htm> (address data 21.02.2021).
209. Mill J.S. A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation. 1843. Cambridge University Press, 2011. 602 p. [Милль Д.С. Система логики силлогистической и индуктивной: Изложение принципов доказательства в связи с методами научного исследования: Пер. с англ. Изд. 5-е, испр. и доп. М.: ЛЕНАНД, 2011, 832 с.]
210. Mackie J.L. Causes and conditions // *Am. Philos. Q.* 1965. V. 2. № 4. P. 245–255.
211. Mackie J.L. The Cement of the Universe: a Study of Causation. Oxford: Clarendon Press, 1974. 329 p.
212. Голосной Д.А. Джон Лесли Мэки о направлении причинения // *Вестн. Моск. ун-та. Сер. 7 “Философия”*. 2014. № 6. С. 30–43. [Golosnoy D.A. John Leslie Mackie about direction of causation // *Vestnik Moskovskogo Universiteta. Seriya 7. “Filosofiya”* (Moscow University Bulletin. Series 7. “Philosophy”). 2014. № 6. P. 30–43. (In Russ. Engl. abstr.)]
213. Hofmann B., Holm S., Iversen J.-G. Philosophy of science // *Research methodology in the medical and biological sciences* / Eds. P. Laake, H.B. Benestad, B.R. Olsen. Academic Press, Elsevier, 2007. P. 1–32.
214. Lewis D. Causation // *J. Philos.* 1973. V. 70. № 17. P. 556–567.
<https://doi.org/10.2307/2025310>
215. Rothman K.J. Causes // *Am. J. Epidemiol.* 1976. V. 104. № 6. P. 587–592.
<https://doi.org/10.1093/oxfordjournals.aje.a112335>
216. Rothman K.J., Poole C. A strengthening programme for weak associations // *Int. J. Epidemiol.* 1988. V. 17. № 4. P. 955–959.
<https://doi.org/10.1093/ije/17.4.955>
217. Rothman K.J. What is causation // *Epidemiology, an introduction* / Ed. K.J. Rothman. New York: Oxford University Press, 2002. P. 8–24.
218. ICRP Publication 118. ICRP Statement on tissue reactions and early and late effects of radiation in normal tissues and organs – threshold doses for tissue reactions in a radiation protection context // *Annals of the ICRP* / Ed. C.H. Clement. Amsterdam–New York: Elsevier, 2012. 325 p.
219. Boice J.D. Jr. The importance of radiation worker studies // *J. Radiol. Prot.* 2014. V. 34. № 3. P. E7–E12.
<https://doi.org/10.1088/0952-4746/34/3/E7>
220. Boice J.D. Jr., Cohen S.S., Mumma M.T., Ellis E.D. The Million Person Study, whence it came and why // *Int. J. Radiat. Biol.* 2019. V. 4. P. 1–14.
<https://doi.org/10.1080/09553002.2019.1589015>
221. UNSCEAR 2000. Report to the General Assembly, with Scientific Annex. Annex H. Combined effects of radiation and other agents. New York, 2000. P. 177–295.
222. Murata M., Miyake T., Inoue Y. et al. Life-style and other characteristics of radiation workers at nuclear facilities in Japan: base-line data of a questionnaire survey // *J. Epidemiol.* 2002. V. 12. № 4. P. 310–319.
<https://doi.org/10.2188/jea.12.310>
223. Telle-Lamberton M., Samson E., Caer S. et al. External radiation exposure and mortality in a cohort of French nuclear workers // *Occup. Environ. Med.* 2007. V. 64. № 10. P. 694–700.
<https://doi.org/10.1136/oem.2007.032631>
224. Vineis P., Veglia F., Benhamou S. et al. CYP1A1 T3801 C polymorphism and lung cancer: a pooled analysis of 2451 cases and 3358 controls // *Int. J. Cancer.* 2003. V. 104. № 5. P. 650–657.
<https://doi.org/10.1002/ijc.10995>
225. Karhausen L.R. Commentary: coda—a Socratic dialogue: Plato // *Int. J. Epidemiol.* 2001. V. 30. № 4. P. 710–711.
<https://doi.org/10.1093/ije/30.4.710>
226. Karhausen L.R. The logic of causation in epidemiology // *Scand. J. Soc. Med.* 1996. V. 24. № 1. P. 8–13.
<https://doi.org/10.1177/140349489602400103>
227. Olsen J. Causes and Prevention // *Scand. J. Soc. Med.* 1991. V. 19. № 1. P. 1–6.
<https://doi.org/10.1177/140349489101900101>
228. Olsen J. Some consequences of adopting a conditional deterministic causal model in epidemiology // *Eur. J. Public Health.* 1993. V. 3. № 3. P. 204–209.
<https://doi.org/10.1093/eurpub/3.3.204>
229. Rothman K.J. Significance questing // *Ann. Intern. Med.* 1986. V. 105. № 3. P. 445–447.
<https://doi.org/10.7326/0003-4819-105-3-445>
230. Rothman K.J., Stuver S.O. Teaching a first course in epidemiologic principles and methods // *Teaching Epidemiology. A Guide for Teachers in Epidemiology, Public Health, and Clinical Medicine*. 4th ed. / Eds J. Olsen, N. Greene, R. Saracci, D. Trichopoulos. New York: Oxford University Press, 2015. P. 67–84.
231. Gatto N.M., Campbell U.B. Redundant causation from a sufficient cause perspective // *Epidemiol. Perspect. Innov.* 2010. V. 7. № 1. Article 5. 13 p.
<https://doi.org/10.1186/1742-5573-7-5>
232. Morris J.N. Uses of epidemiology. 1955 // *Int. J. Epidemiol.* 2007. V. 36. № 6. P. 1165–1172.
<https://doi.org/10.1093/ije/dym227>
233. Morris J.N. Uses of Epidemiology. Edinburgh and London: E. & S. Livingstone, 1957. 135 p.
234. Davey Smith G. The uses of “Uses of Epidemiology” // *Int. J. Epidemiol.* 2001. V. 30. № 5. P. 1146–1155.
<https://doi.org/10.1093/ije/30.5.1146>
235. Morris J.N. Modern epidemiology? (Letter to the editor) // *J. Epidemiol. Community Health.* 1988. V. 42. № 1. P. 100.

236. *Davey Smith G.* Epidemiology, epigenetics and the “Gloomy Prospect”: embracing randomness in population health research and practice // *Int. J. Epidemiol.* 2011. V. 40. № 3. P. 537–562. <https://doi.org/10.1093/ije/dyr117>
237. *Senn S.* Statistical pitfalls of personalized medicine // *Nature.* 2018. V. 563. № 7733. P. 619–621. <https://doi.org/10.1038/d41586-018-07535-2>
238. *Laplace P.S.* A Philosophical Essay on Probabilities. New York: Dover Publications, 1951. 204 p. [*Ланлас П.С.* Опыт философии теории вероятностей. М.: Книжный дом “ЛИБРОКОМ”, 2011. 208 с. (репринт издания 1814 г.)]
239. *Fisher R.A.* Indeterminism and natural selection // *Philosophy of Science.* 1934. V. 1. № 1. P. 99–117. <https://doi.org/10.1086/286308.hdl:2440/15119>
240. *Sturmberg J.P.* Evidence-based medicine – not a panacea for the problems of a complex adaptive world // *J. Eval. Clin. Pract.* 2019. V. 25. № 5. P. 706–716. <https://doi.org/10.1111/jep.13122>
241. *Greenland S., Morgenstern H.* Confounding in health research // *Annu. Rev. Public Health.* 2001. V. 22. P. 189–212. <https://doi.org/10.1146/annurev.publhealth.22.1.189>
242. *Hill L.G., Rosenman R., Tennekoon V., Mandal B.* Selection effects and prevention program outcomes // *Prev. Sci.* 2013. V. 14. № 6. P. 557–569. <https://doi.org/10.1007/s1121-012-0342-x>
243. *Stock J.H., Watson M.W.* Introduction to Econometrics. 3rd ed. Addison-Wesley US, 2011. 785 p. [*Сток Д., Уотсон М.* Введение в эконометрику: Пер. с англ. / Под ред. М.Ю. Турунцевой. М.: Издательский дом “Дело” РАНХиГС, 2015 836 с.]
244. *Maldonado G., Greenland S.* Estimating causal effects // *Int. J. Epidemiol.* 2002. V. 31. № 2. P. 422–429. <https://doi.org/10.1093/intjepid/31.2.422>
245. *Iwami O., Ikeda M.* Steady state trials: another valid substitution of counterfactual ideal to measure causal effects // *Environ. Health Prev. Med.* 2013. V. 18. № 3. P. 237–243. <https://doi.org/10.1007/s12199-012-0312-8>
246. *Handbook of Epidemiology.* 2nd ed. / Eds W. Ahrens, I. Pigeot. New York, Heidelberg, Dordrecht, London: Springer, 2014. 2498 p.
247. *Weisberg H.I., Hayden V.C., Pontes V.P.* Selection criteria and generalizability within the counterfactual framework: explaining the paradox of antidepressant-induced suicidality? // *Clin. Trials.* 2009. V. 6. № 2. P. 109–118. <https://doi.org/10.1177/1740774509102563>
248. *McGue M., Osler M., Christensen K.* Causal inference and observational research: the utility of twins // *Perspect. Psychol. Sci.* 2010. V. 5. № 5. P. 546–556. <https://doi.org/10.1177/1745691610383511>
249. *Hart S.A., Taylor J., Schatschneider C.* There is a world outside of experimental designs: using twins to investigate causation // *Assess Eff. Interv.* 2013. V. 38. № 2. P. 117–126. <https://doi.org/10.1177/1534508412451490>
250. *Burt S.A., Donnellan M.B., Humbad M.N. et al.* Does marriage inhibit antisocial behavior? An examination of selection vs causation via a longitudinal twin design // *Arch. Gen. Psychiatry.* 2010. V. 67. № 12. P. 1309–1315. <https://doi.org/10.1001/archgenpsychiatry.2010.159>
251. *Flanders W.D., Eldridge R.C.* Summary of relationships between exchangeability, biasing paths and bias // *Eur. J. Epidemiol.* 2015. V. 30. № 10. P. 1089–1099. <https://doi.org/10.1007/s10654-014-9915-2>
252. *Lu M., Sadiq S., Feaster D.J., Ishwaran H.* Estimating individual treatment effect in observational data using random forest methods // *J. Comput. Graph. Stat.* 2018. V. 27. № 1. P. 209–219. <https://doi.org/10.1080/10618600.2017.1356325>
253. *Holland P.W.* Counterfactual reasoning // *Encyclopedia of Statistics in Behavioral Science.* V. 1 / Eds B.S. Everit, D.C. Howell. Chichester: John Wiley & Sons, Ltd, 2005. 420–422.
254. *Hume D.* An Enquiry Concerning Human Understanding. Open Court Press, LaSalle, 1748 (reprint 1888). [*Юм Д.* Сочинения в 2 т. Т. 2: Пер. с англ. С.И. Церетели и др. 2-е изд., дополн. и испр. М.: Мысль, 1996. 733 с. (цитата о контрафактической причинности: с. 65–66)]
255. *Rubin D.B.* Comment: Neyman (1923) and causal inference in experiments and observational studies // *Stat. Sci.* 1990. V. 5. № 4. P. 472–480.
256. *Neyman J.* (1923/1990) Sur les applications de la theorie des probabilites aux experiences agricoles [On the application of probability theory to agricultural experiments: essay on principles (Section 9)] / Translation and edition by D. Dabrowska, T.P. Speed // *Stat. Sci.* 1990. V. 5. № 4. P. 465–472. <https://doi.org/10.1214/ss/1177012031>, MR 1092986
257. *Rubin D.B.* Estimating causal effects of treatments in randomized and nonrandomized studies // *J. Educat. Psychol.* 1974. V. 66. № 5. P. 688–701. <https://doi.org/10.1037/h0037350>
258. *Greenland S., Robins J.* Identifiability, exchangeability, and epidemiological confounding // *Int. J. Epidemiol.* 1986. V. 15. № 3. P. 413–419. <https://doi.org/10.1093/ije/15.3.413>
259. *Greenland S., Pearl J., Robins J.M.* Confounding and collapsibility in causal inference // *Stat. Sci.* 1999. V. 14. № 1. P. 29–46. <https://doi.org/10.1214/ss/1009211805>
260. *Hogan J.W.* Bringing causal models into the mainstream // *Epidemiology.* 2009. V. 20. № 3. P. 431–432. <https://doi.org/10.1097/EDE.0b013e3181a0997a>
261. *Levine B.* Causal models. Letters to the Editor // *Epidemiology.* 2009. V. 20. № 6. P. 931. <https://doi.org/10.1097/EDE.0b013e3181ba451f>
262. *Huang Y.T., Tai A.S., Chou M.Y. et al.* Six-way decomposition of causal effects: Unifying mediation and mechanistic interaction // *Stat. Med.* 2020. V. 39. № 27. P. 4051–4068. <https://doi.org/10.1002/sim.8708>
263. *Shafer G.* The Art of Causal Conjecture. Cambridge: MIT Press, 1996. 531 p.
264. *Dawid A.P.* Causal inference without counterfactuals // *J. Am. Stat. Assoc.* 2000. V. 95. № 450. P. 407–424. <https://doi.org/10.1080/01621459.2000.10474210>
265. *Doran C.M., Ling R., Byrnes J.* Benefit cost analysis of three skin cancer public education mass-media campaigns implemented in New South Wales, Australia //

- PLoS One. 2016. V. 11. № 1. Article e0147665. <https://doi.org/10.1371/journal.pone.0147665>
266. Olsen C.M., Wilson L.F., Green A.C. et al. How many melanomas might be prevented if more people applied sunscreen regularly? // *Br. J. Dermatol.* 2018. V. 178. № 1. P. 140–147. <https://doi.org/10.1111/bjd.16079>
267. De Vocht F. Inferring the 1985–2014 impact of mobile phone use on selected brain cancer subtypes using Bayesian structural time series and synthetic controls // *Environ. Int.* 2016. V. 97. P. 100–107. <https://doi.org/10.1016/j.envint.2016.10.019>
268. Sterne J. An Introduction to causal inference. Presentation at the 2018 Society for Research Synthesis Methodology meeting. Bristol, UK, July 17–19, 2018.
269. Greenland S., Pearl J., Robins J.M. Causal diagrams for epidemiologic research // *Epidemiology.* 1999. V. 10. № 1. P. 37–48. <https://doi.org/10.1097/00001648-199901000-00008>
270. Suttorp M.M., Siegerink B., Jager K.J. et al. Graphical presentation of confounding in directed acyclic graphs // *Nephrol. Dial Transplant.* 2015. V. 30. № 9. P. 1418–1423. <https://doi.org/10.1093/ndt/gfu325>
271. Tennant P.W.G., Harrison W.J., Murray E.J. et al. Use of directed acyclic graphs (DAGs) in applied health research: review and recommendations // *Int. J. Epidemiol.* 2020. Article dyaa213. 135 p. <https://doi.org/10.1093/ije/dyaa213>
272. Yamamoto W.S., Wolff P.G. On the identification of verbs in computer programs of physiological models // *Comput. Biomed. Res.* 1984. V. 17. № 2. P. 175–184. [https://doi.org/10.1016/0010-4809\(84\)90030-2](https://doi.org/10.1016/0010-4809(84)90030-2)
273. Тулупьев А.Л., Сироткин А.В., Николенко С.И. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.: Изд-во СПб. ун-та, 2009. 400 с. [Tulupuev A.L., Sirotkin A.V., Nikolenko S.I. Bayesian Trust Networks: Probabilistic Inference in Acyclic Directed Graphs. Saint Petersburg: Publishing house of Saint Petersburg University, 2009. 400 p. (In Russ.)]
274. Шехтер Д.Б., Чадюк А.В., Червинский-Ивашура А.Л. Байесовские методы в задаче оценки релевантности при поиске работы в Internet // Проблемы программирования. 2006. № 2-3. Спец. вып. С. 519–525. [Shekhter D.B., Chadyuk A.V., Chervinsky-Ivashura A.L. Bayesian methods in the problem of assessing the relevance when looking for a job on the Internet // Problems of programming. 2006. № 2–3. Special issue. P. 519–525. (In Russ. Engl. abstr.)]
275. Торопова А.В. Байесовские сети доверия: инструменты и использование в учебном процессе // Компьютерные инструменты в образовании. 2016. № 4. С. 43–53. [Toropova A.V. Bayesian belief networks: a tool for education use // Computer Tools in Education. 2016. № 4. P. 43–53. (In Russ. Engl. abstr.)]
276. Золотин А.А. Развитие логико-вероятностного вывода в алгебраических байесовских сетях // Компьютерные инструменты в образовании. 2017. № 3. С. 5–19. [Zolotin A.A. Development of probabilistic-logic inference in algebraic Bayesian networks // Computer Tools in Education. 2017. № 3. P. 5–19. (In Russ. Engl. abstr.)]
277. Димитрова Л.К., Голубева О.А. Применение байесовской сети в дифференциальной диагностике артериальной гипертензии // Технические науки в России и за рубежом: Матер. III Междунар. науч. конф. Москва, июль 2014 г. М.: Буки-Веди, 2014. С. 4–14. [Dimitrova L.K., Golubeva O.A. Application of Bayesian network in differential diagnosis of arterial hypertension // Technical sciences in Russia and abroad: Mater. III Intern. scientific conf. Moscow, July 2014. Moscow: Buki-Vedi, 2014. P. 4–14. (In Russ.)]
278. Kharyuk P., Nazarenko D., Oseledets I. Employing fingerprinting of medicinal plants by means of LC-MS and machine learning for species identification task // *Sci. Rep.* 2018. V. 8. № 1. Article 17053. <https://doi.org/10.1038/s41598-018-35399-z>
279. Самоходская Л.М., Старостина Е.Е., Сулимов А.В. и др. Прогнозирование особенностей течения хронического гепатита С с использованием байесовских сетей // Тер. архив. 2019. Т. 91. № 2. С. 32–39. [Samokhodskaya L.M., Starostina E.E., Sulimov A.V. et al. Prediction of features of the course of chronic hepatitis C using Bayesian networks // Ter. Arkh. 2019. V. 91. № 2. P. 32–39. (In Russ.; Engl.)] <https://doi.org/10.26442/00403660.2019.02.000076>
280. Greenland S. Causation and Causal Inference // *International Encyclopedia of Statistical Science* / Ed. M. Lovric. 2011. P. 216–221. https://doi.org/10.1007/978-3-642-04898-2_9
281. Basu A. Directed acyclic graphs to explore causality in epidemiological study designs, part I: an introduction to DAGs // *Qeios.* 2020. <https://doi.org/10.32388/FFH3GU>. <https://www.qeios.com/read/FFH3GU>.
282. Handbook of Causal Analysis for Social Research / Ed. S.L. Morgan. Dordrecht etc.: Springer, 2013. 424 p.
283. Pearl J. An Introduction to Causal Inference. The Round Thing, 2014. 229 p.
284. Foraita R., Spallek J., Zeeb H. Directed Acyclic Graphs // *Handbook of Epidemiology.* 2nd ed. / Eds W. Ahrens, I. Pigeot. New York, Heidelberg, Dordrecht, London: Springer, 2014. P. 1481–1617.
285. Robins J.M. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods // *J. Chron. Dis.* 1987. V. 40. Suppl. 2. P. 139s–161s. [https://doi.org/10.1016/S0021-9681\(87\)80018-8](https://doi.org/10.1016/S0021-9681(87)80018-8)
286. Williams T.C., Bach C.C., Matthesen N.B. et al. Directed acyclic graphs: a tool for causal studies in paediatrics // *Review Pediatr. Res.* 2018. V. 84. № 4. P. 487–493. <https://doi.org/10.1038/s41390-018-0071-3>
287. Николенко С. Байесовские сети доверия. Ч. I и II. Самообучающиеся системы (Machine Learning). Нейронные сети машинное обучение. Лекции 8 и 9 (презентации) // CSIN-Ru. Информатика в России, 2006. [Nikolenko S. Bayesian belief networks. Part I and II. Self-learning systems (Machine Learning). Neural networks machine learning. Lektion 8 and 9 (presentations) // CSIN-Ru, Informatics in Russia, 2006. (In Russ.)] <https://web.archive.org/web/20071016161230/http://www.csin.ru/c>

- ourses/samoobuchayushchiesya-sistemy-i-nechetkaya-logika (address data 05.03.2021).
288. *Reddan T., Corness J., Harden F. et al.* Bayesian approach to predicting acute appendicitis using ultrasonographic and clinical variables in children // *Healthc. Inform. Res.* 2019. V. 25. № 3. P. 212–220. <https://doi.org/10.4258/hir.2019.25.3.212>
 289. *Morabia A.* Mervyn Susser, the last of the three American classical epidemiology tenors // *Ann. Epidemiol.* 2015. V. 25. № 2. P. 140–142. <https://doi.org/10.1016/j.annepidem.2014.10.014>
 290. *Susser M.* Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology. New York: Oxford Univ. Press, 1973. 181 p.
 291. *Blalock H.M.* Causal inference in non-experimental research. Chapel Hill, NC: University of North Carolina Press, 1964. 200 p.
 292. *Pearl J.* Probabilistic Reasoning in Intelligent Systems. San Mateo, CA: Morgan Kaufmann, 1988. 576 p.
 293. *Greenland S., Neutra R.* Control of confounding in the assessment of medical technology // *Int. J. Epidemiol.* 1980. V. 9. № 4. P. 361–367. <https://doi.org/10.1093/ije/9.4.361>
 294. *VanderWeele T.J., Hernan M.A., Robins J.M.* Causal directed acyclic graphs and the direction of unmeasured confounding bias // *Epidemiology.* 2008. V. 19. № 5. P. 720–728. <https://doi.org/10.1097/EDE.0b013e3181810e29>
 295. *Methods in Social Epidemiology.* 2nd ed. / Eds J.M. Oakes, J.S. Kaufman. San Francisco: Jossey-Bass. A Wiley Brand, 2017. 568 p.
 296. *Long R.* Reverse causality opposite definitions // Site “Cross Validated”. Sponsored by “AWS Machine learning”. Stack Exchange Inc., 2021. <https://stats.stackexchange.com/questions/445227/reverse-causality-opposite-definitions> (address data 05.03.2021).
 297. *Котеров А.Н., Ушенкова Л.Н., Калинина М.В., Бирюков А.П.* Краткий обзор мировых исследований лучевых и нелучевых эффектов у работников ядерной индустрии // *Медико-биологические проблемы жизнедеятельности (Гомель).* 2020. № 1. С. 17–31. [*Kotero A.N., Ushenkova L.N., Kalinina M.V., Biryukov A.P.* Brief review of world researches of radiation and non-radiation effects in nuclear industry workers // *Medical and Biological Problems of Life Activity (Gomel).* 2020. № 1. P. 17–31. (In Russ. Engl. abstr.)]
 298. *Yule G.U.* On the theory of correlation for any number of variables // *Proc. Roy. Soc. (A).* 1907. V. 79. P. 182–193.
 299. *Горбачева И.* Краткий курс по методам математической статистики. Сайт “Медицинская статистика”. 2013. 43 с. [*Gorbacheva I.* Short course on methods of mathematical statistics. Website “Medical statistics”. 2013. 43 p. (In Russ.)] http://medstatistic.ru/articles/kratkiy_kurs.pdf (address data 27.04.2021).
 300. *Denis D.J.* Applied Univariate, Bivariate, and Multivariate Statistics. Hoboken, New Jersey. John Wiley & Sons, Inc., 2016. 726 p.
 301. *Audino P., La Grutta S., Cibella F. et al.* Rhinitis as a risk factor for depressive mood in pre-adolescents: a new approach to this relationship // *Pediatr. Allergy Immunol.* 2014. V. 25. № 4. P. 360–365. <https://doi.org/10.1111/pai.12215>
 302. *Cernis E., Evans R., Ehlers A., Freeman D.* Dissociation in relation to other mental health conditions: An exploration using network analysis // *J. Psychiatr. Res.* 2021. V. 136. P. 460–467. <https://doi.org/10.1016/j.jpsychires.2020.08.023>
 303. *Ogpen-Rhein R., Strimmer K.* From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data // *BMC Syst. Biol.* 2007. V. 1. Article 37. 10 p. <https://doi.org/10.1186/1752-0509-1-37>
 304. *Freudenberg J., Wang M., Yang Y., Li W.* Partial correlation analysis indicates causal relationships between GC-content, exon density and recombination rate in the human genome // *BMC Bioinformatics.* 2009. V. 10. Suppl. 1. Art. 566. 10 p. <https://doi.org/10.1186/1471-2105-10-S1-S66>
 305. *Shuvalova Y.A., Khasanova Z.B., Kaminnaya V.I. et al.* The association of PLA2G2A single nucleotide polymorphisms with type IIa secretory phospholipase A2 level but not its activity in patients with stable coronary heart disease // *Gene.* 2015. V. 564. № 1. P. 29–34. <https://doi.org/10.1016/j.gene.2015.03.030>
 306. *Yazdani A., Yazdani A., Samiei A., Boerwinkle E.* Generating a robust statistical causal structure over 13 cardiovascular disease risk factors using genomics data // *J. Biomed. Inform.* 2016. V. 60. P. 114–119. <https://doi.org/10.1016/j.jbi.2016.01.012>
 307. *Ha M.J., Sun W.* Estimation of high-dimensional directed acyclic graphs with surrogate intervention // *Biostatistics.* 2020. V. 21. № 4. P. 659–675. <https://doi.org/10.1093/biostatistics/kxy080>
 308. *Michna A., Schotz U., Selmansberger M. et al.* Transcriptomic analyses of the radiation response in head and neck squamous cell carcinoma subclones with different radiation sensitivity: time-course gene expression profiles and gene association networks // *Radiat. Oncol.* 2016. V. 11. Article 94. 16 p. <https://doi.org/10.1186/s13014-016-0672-0>
 309. *Alexopoulou E., Katsila T., Tolia M. et al.* An exploratory study of radiation dermatitis in breast cancer patients // *Anticancer Res.* 2018. V. 38. № 3. P. 1615–1622. <https://doi.org/10.21873/anticancer.12392>
 310. *Viel J.F.* Radon exposure and leukaemia in adulthood // *Int. J. Epidemiol.* 1993. V. 22. № 4. P. 627–631. <https://doi.org/10.1093/ije/22.4.627>
 311. *Chen R., Seaton A.* The influence of study characteristics on the healthy worker effect: a multiple regression analysis // *Occup. Med. (Lond).* 1996. V. 46. № 5. P. 345–350. <https://doi.org/10.1093/occmed/46.5.345>
 312. *Walsh L.* Neutron relative biological effectiveness for solid cancer incidence in the Japanese A-bomb survivors: an analysis considering the degree of independent effects from γ -ray and neutron absorbed doses with hierarchical partitioning // *Radiat. Environ. Biophys.* 2013. V. 52. № 1. P. 29–36. <https://doi.org/10.1007/s00411-012-0445-6>

313. Lee J.K., Lee M.S., Moon M.H. et al. Translocation frequency in patients with repeated CT exposure: Comparison with CT-naive patients // *Radiat. Res.* 2019. V. 192. № 1. P. 23–27. <https://doi.org/10.1667/RR15286.1>
314. Boratynski Z., Arias J.M., Garcia C. et al. Ionizing radiation from Chernobyl affects development of wild carrot plants // *Sci. Rep.* 2016. V. 6. Article. 39282. 8 p. <https://doi.org/10.1038/srep39282>
315. Gyuleva I., Djounova J., Rupova I. Impact of low-dose occupational exposure to ionizing radiation on T-cell populations and subpopulations and humoral factors included in the immune response // *Dose Response.* 2018. V. 16. № 3. Article 1559325818785564. 8 p. <https://doi.org/10.1177/1559325818785564>
316. Belyaev I.Ya., Spivak I.M., Kolman A., Harms-Ringdahl M. Relationship between radiation induced adaptive response in human fibroblasts and changes in chromatin conformation // *Mutat. Res.* 1996. V. 358. № 2. P. 223–230. [https://doi.org/10.1016/s0027-5107\(96\)00124-8](https://doi.org/10.1016/s0027-5107(96)00124-8)
317. Saberi A.H., Kee B.B., Oi-Ming L., Miskandar M.S. Physico-chemical properties of various palm-based diacylglycerol oils in comparison with their corresponding palm-based oils // *Food Chem.* 2011. V. 127. № 3. P. 1031–1038. <https://doi.org/10.1016/j.foodchem.2011.01.076>
318. Beral V., Fraser P., Carpenter L. et al. Mortality of employees of the Atomic Weapons Establishment, 1951–82 // *Brit. Med. J.* 1988. V. 297. № 6651. P. 757–770. <https://doi.org/10.1136/bmj.297.6651.757>
319. Atkinson W.D., Law D.V., Bromley K.J., Inskip H.M. Mortality of employees of the United Kingdom Atomic Energy Authority, 1946–97 // *Occup. Environ. Med.* 2004. V. 61. № 7. P. 577–585. <https://doi.org/10.1136/oem.2003.012443>
320. Gilbert E.S., Omohundro E., Buchanan J.A., Holter N.A. Mortality of workers at the Hanford site: 1945–1986 // *Health Phys.* 1993. V. 64. № 6. P. 577–590. <https://doi.org/10.1097/00004032-199306000-00001>
321. Carpenter L.M., Higgins C.D., Douglas A.J. et al. Cancer mortality in relation to monitoring for radionuclide exposure in three UK nuclear industry workforces // *Br. J. Cancer.* 1998. V. 78. № 9. P. 1224–1232. <https://doi.org/10.1038/bjc.1998.659>
322. Muirhead C.R., O'Hagan J.A., Haylock R.G.E. et al. Mortality and cancer incidence following occupational radiation exposure: third analysis of the National Registry for Radiation Workers // *Brit. J. Cancer.* 2009. V. 100. № 1. P. 206–212. <https://doi.org/10.1038/sj.bjc.6604825>
323. Gillies M., Haylock R. The cancer mortality and incidence experience of workers at British Nuclear Fuels plc, 1946–2005 // *J. Radiol. Prot.* 2014. V. 34. № 3. P. 595–623. <https://doi.org/10.1088/0952-4746/34/3/595>
324. Котеров А.Н., Туков А.Р., Ушенкова Л.Н. и др. Средняя накопленная доза облучения для работников мировой ядерной индустрии: малые дозы, малые эффекты. Сравнение с дозами для медицинских радиологов // *Мед. радиология и радиац. безопасность.* 2021. Т. 66. № 2. [Kotero A.N., Tukov A.R., Ushenkova L.N. et al. Average accumulated radiation doses for world nuclear workers: low doses, low effects. Comparison with doses for medical radiologists // *Medits. Radiologija Radiat. Bezopasnost* (“Medical Radiology and Radiation Safety”; Moscow). 2021. V. 66. № 2. (In Russ. Engl. abstr.)]
325. DOE 1995. U.S. Department of Energy. Closing the Circle on the Splitting of the Atom // *The Environmental Legacy of Nuclear Weapons Production in the United States and What the Department of Energy is Doing About It.* U.S. Department of Energy, Office of Environmental Management, January 1995. DOE/EM-0266. 106 p. https://www.energy.gov/sites/prod/files/2014/03/f8/Closing_the_Circle_Report.pdf (address data 27.04.2021).
326. Muirhead C.R., O'Hagan J.A., Haylock R.G.E. et al. Third analysis of the National Registry for Radiation Workers: occupational exposure to ionizing radiation in relation to mortality and cancer incidence. Health Protection Agency. Centre for Radiation, Chemical and Environmental Hazards. Radiation Protection Division. HPA-RPD-062. Chilton, Didcot, Oxfordshire OX11 0RQ, 2009. 150 p.
327. Deltour I., Tretyakov F., Tsareva Y. et al. Mortality of populations potentially exposed to ionising radiation, 1953–2010, in the closed city of Ozyorsk, Southern Urals: a descriptive study // *Environ. Health.* 2015. V. 14. Art. 91. 12 p. <https://doi.org/10.1186/s12940-015-0078-8>
328. Кошурникова Н.А., Окатенко П.В., Сокольников М.Э. и др. Медицинские последствия профессионального облучения: канцерогенный риск в когорте персонала ПО “МАЯК” // *Мед. радиология и радиац. безопасность.* 2008. Т. 53. № 3. С. 23–33. [Koshurnikova N.A., Okatenko P.V., Sokolnikov M.E. et al. Medical consequences of the professional exposure: carcinogenic risk in the cohort of “Mayak” PA workers // *Medits. Radiologija Radiat. Bezopasnost* (“Medical Radiology and Radiation Safety”; Moscow). 2008. V. 53. № 3. P. 23–33. (In Russ. Engl. abstr.)]
329. Жунтова Г.В., Азизова Т.В., Банникова М.В., Фомин Е.П. Структура причин смерти работников предприятия атомной промышленности, подвергшихся профессиональному облучению в различном диапазоне доз // *Вопр. радиац. безопасности.* 2017. № 2 (86). С. 47–57. [Zhuntova G.V., Azizova T.V., Bannikova M.V., Fomin E.P. The structure of the causes of death of nuclear industry workers exposed to occupational exposure in various dose ranges // *Issues of radiation safety.* 2017. № 2 (86). P. 47–57. (In Russ. Engl. abstr.)]
330. Cancer Research UK. Cancer mortality by age. Appendix “Deaths_crude_mf_allcancer_m18.xlsx”. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/age#heading-One> (address data 27.04.2021).
331. White M.C., Holman D.M., Boehm J.E. et al. Age and cancer risk: a potentially modifiable relationship // *Am. J. Prev. Med.* 2014. V. 46. № 3 (Suppl. 1. P. S7–S15. <https://doi.org/10.1016/j.amepre.2013.10.029>
332. Liu L. Global sex differences in cancer mortality with age and country specific characteristics // *Asian Pac. J. Cancer Prev.* 2016. V. 17. № 7. P. 3469–3476.

333. Gu X., Zheng R., Xia C. *et al.* Interactions between life expectancy and the incidence and mortality rates of cancer in China: a population-based cluster analysis // *Cancer Commun. (Lond)*. 2018. V. 38. № 1. Article 44. 15 p.
https://doi.org/10.1186/s40880-018-0308-x
334. Boice J.D. Jr. Models, models everywhere – is there a fit for lifetime risks? (Editorial) // *J. Natl. Cancer Inst.* 2010. V. 102. № 21. P. 1606–1609.
https://doi.org/10.1093/jnci/djq412
335. Krieger N., Smith G.D. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology // *Int. J. Epidemiol.* 2016. V. 45. № 6. P. 1787–808.
https://doi.org/10.1093/ije/dyw114
336. Geneletti S., Gallo V., Porta M. *et al.* Assessing causal relationships in genomics: From Bradford-Hill criteria to complex gene-environment interactions and directed acyclic graphs // *Emerg. Themes Epidemiol.* 2011. V. 8. № 1. Art. 5. 18 p.
https://doi.org/10.1186/1742-7622-8-5
337. Wu X. Efficient intensity map splitting algorithms for intensity-modulated radiation therapy // *Inf. Process Lett.* 2008. V. 106. № 5. P. 188–194.
https://doi.org/10.1016/j.ipl.2007.11.008
338. Wu X., Dou X., Bayouth J., Buatti J.M. An almost linear time algorithm for field splitting in radiation therapy // *Comput. Geom.* 2013. V. 46. № 6. P. 673–687.
https://doi.org/10.1016/j.comgeo.2012.11.001
339. Lee S., Ybarra N., Jeyaseelan K. *et al.* Bayesian network ensemble as a multivariate strategy to predict radiation pneumonitis risk // *Med. Phys.* 2015. V. 42. № 5. P. 2421–2430.
https://doi.org/10.1118/1.4915284
340. Kroenke C.H., Neugebauer R., Meyerhardt J. *et al.* Analysis of body mass index and mortality in patients with colorectal cancer using causal diagrams // *J. Am. Med. Assoc. Oncol.* 2016. V. 2. № 9. P. 1137–1145.
https://doi.org/10.1001/jamaoncol.2016.0732
341. Stewart-Oaten A. Problems in the analysis of environmental monitoring data // *Detecting Ecological Impacts: Concepts and Applications in Coastal Habitats* / Eds R.J. Schmitt, C.W. Osenberg. San Diego, CA: Academic Press, 1996. P. 109–131.

Causal Criteria in Medical and Biological Disciplines: History, Essence and Radiation Aspect. Report 4, Part 2: Hierarchy of Criteria, Their Criticism and other Methods for Causation Establishing

A. N. Koterov^{a, #} and L. N. Ushenkova^a

^a A.I. Burnasyan Federal Medical Biophysical Center of Federal Medical Biological Agency, Moscow, Russia

[#] E-mail: govorilga@inbox.ru

Part 2 of Report 4 covers three topics. First, attempts to systematize, weighing and form a hierarchy (“ranks”) for the causal criteria. Two methodologies were used: referring to specific constructions of various researchers and assessment by determining the frequency of occurrence of a particular criterion in those works in which they were used as a way to establish the effect causality. Apart from the *a priori* leading criterion “Temporality”, the first two places should be given to “Strength of association” and “Consistency of association” (their sequence may be reversed), that is, inductive points. The third place is taken by the “Counterfactual experiment” (“natural experiment”), which, however, is rarely available in epidemiology and ecology. In its absence, the third place is again taken by the inductive criterion “Biological gradient” (relationship “dose–effect”). In the fields of ecology and ecotoxicology, where experimental or observational studies in humans are rarely available, the ranks of the criteria are specific, and “Biological plausibility” is in the first place. Second, it presents a criticism of the causal approach in observational disciplines based on criteria of causality, starting from the lack of evidence of the inductive method itself. It points to the “mosaicism”, multidisciplinary and subjectivity of the criteria, as well as their non-absoluteness (“are saddled with reservations and exceptions”). It is noted, however, that such questions can be posed before almost any approach in the field of natural sciences. There is an opinion that the criteria of causality, as it were “confirming effects”, should be left only for decision-making in the field of health care, and not for strict evidence in the field of “pure science” with its falsification of hypotheses. It is unlikely that this approach can be adopted for epidemiology.

Third, models for assessing the causality of effects in epidemiology outside of causal criteria are considered. The most famous is the K.J. Rothman (“Causality Pie”), implying determinism of epidemiological effects, as opposed to the position of practical epidemiology and public health, in which causality is defined as probabilistic. The second model is called the “counterfactual” or “potential outcome model”, the third is directed acyclic graphs or causal diagrams (DAG), and the fourth is the “Structural equation model”. Another approach was discovered by C.V. Phillips, which quantifies errors (uncertainties) at all stages of an epidemiological study with Monte Carlo simulation. It is concluded that the named models turn out to be either qualitative and illustrative, or represent only a graphic or mathematical apparatus that still serves the same initial approach based on causal criteria.

Keywords: causal criteria, criteria of A.B. Hill, sufficient component causal model, counterfactual model, directed acyclic graph model, structural equation model, epidemiology