



Базовые аспекты мета-анализа. Часть 1

А.Ю. Суворов, И.В. Латушкина[✉], К.А. Гуляева, Н.М. Буланов, М.Ю. Надинская, А.А. Заикин

ФГАОУ ВО «Первый Московский государственный медицинский университет им. И.М. Сеченова»
Минздрава России (Сеченовский Университет)
ул. Трубецкая, д. 8, стр. 2, г. Москва, 119991, Россия

Аннотация

Мета-анализ – одно из понятий научной методологии. Он является частым, но не обязательным компонентом систематического обзора эмпирических исследований. Для проведения мета-анализа объединяются результаты нескольких научных исследований и осуществляется проверка одной или нескольких взаимосвязанных научных гипотез при помощи количественных (статистических) методов. Для такого анализа можно использовать либо первичные данные оригинальных исследований, либо обобщенные опубликованные (вторичные) результаты исследований, посвященные одной проблеме. Мета-анализ используется для получения оценки величины неизвестного эффекта, а также для сравнения результатов различных исследований, выявляет в них закономерности или другие взаимосвязи, а также возможные источники разногласий. Мета-анализы занимают высшую ступень достоверности в концепции доказательной медицины, поэтому их результаты считаются самым надежным источником доказательств. Понимание всех этапов проведения мета-анализа позволит научным сотрудникам грамотно анализировать результаты таких исследований, а также формулировать задачи при самостоятельном проведении мета-анализов. В настоящей статье читатель познакомится с такими ключевыми понятиями мета-анализа, как взвешенные эффекты, гетерогенность, различные типы используемых статистических моделей, а также научится работать с некоторыми видами графиков, получаемых в мета-анализах.

Ключевые слова: размер эффекта; модель с фиксированными эффектами; модель со случайными эффектами; гетерогенность; анализ чувствительности; рандомизированное контролируемое исследование; когортное исследование

Рубрики MeSH:

МЕТА-АНАЛИЗ КАК ТЕМА

Для цитирования: Суворов А.Ю., Латушкина И.В., Гуляева К.А., Буланов Н.М., Надинская М.Ю., Заикин А.А. Базовые аспекты мета-анализа. Часть 1. Сеченовский вестник. 2023; 14(1): 4–14. <https://doi.org/10.47093/2218-7332.2023.14.1.4-14>

КОНТАКТНАЯ ИНФОРМАЦИЯ:

Латушкина Ирина Викторовна, младший научный сотрудник Центра анализа сложных систем ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет)

Адрес: ул. Трубецкая, д. 8, стр. 2, г. Москва, 119991, Россия

Тел.: +7 (916) 126-12-85

E-mail: latushkina_i_v@staff.sechenov.ru

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Статья подготовлена при поддержке программы стратегического академического лидерства «Приоритет-2030» ФГАОУ ВО «Первый Московский государственный медицинский университет им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

Поступила: 22.12.2022

Принята: 09.02.2023

Дата печати: 30.03.2023

Basic aspects of meta-analysis. Part 1

Alexander Yu. Suvorov, Irina V. Latushkina✉, Kseniya A. Gulyaeva, Nikolay M. Bulanov,
Maria Yu. Nadinskaia, Alexey A. Zaikin

*Sechenov First Moscow State Medical University (Sechenov University)
8/2, Trubetskaya str., Moscow, 119991, Russia*

Abstract

Meta-analysis is one of the concepts of scientific methodology, and is a frequent but optional component of systematic reviews of empirical research. It joins the results of several scientific studies and tests one or more interrelated scientific hypotheses using quantitative (statistical) methods. This analysis can either use primary data from the original studies or published (secondary) results of studies dealing with the same problem. Meta-analysis is used to obtain an estimate of the magnitude of an unknown effect, and compare the results of different studies, identifying patterns or other relationships in them, as well as possible sources of disagreement. Meta-analyses are the highest level of credibility within evidence-based medicine (EBM), so meta-analysis results are considered as the most reliable source of evidence. Understanding all the procedures of a meta-analysis will allow researchers to analyze the results of such studies correctly, as well as formulate tasks when conducting meta-analyses on their own. In this article the reader will be introduced to key concepts such as weighted effects, heterogeneity, the different types of statistical models used, and how to work with some of the types of plots produced in meta-analyses.

Keywords: effect size; fixed effects model; random effects model; heterogeneity; sensitivity analysis; randomized controlled trial; cohort study

MeSH terms:

META-ANALYSIS AS TOPIC

For citation: Suvorov A.Yu., Latushkina I.V., Gulyaeva K.A., Bulanov N.M., Nadinskaia M.Yu., Zaikin A.A. Basic aspects of meta-analysis. Part 1. Sechenov Medical Journal. 2023; 14(1): 4–14. <https://doi.org/10.47093/2218-7332.2023.14.1.4-14>

CONTACT INFORMATION:

Irina V. Latushkina, junior researcher, Centre for Analysis of Complex Systems, Sechenov First Moscow State Medical University (Sechenov University)

Address: 8/2, Trubetskaya str., Moscow, 119991, Russia

Tel.: +7 (916) 126-12-85

E-mail: latushkina_i_v@staff.sechenov.ru

Conflict of interests. The authors declare that there is no conflict of interests.

Financial support. This article was supported by the Academic leadership program Priority 2030 proposed by Sechenov First Moscow State Medical University (Sechenov University)

Received: 22.12.2022

Accepted: 09.02.2023

Date of publication: 30.03.2023

Список сокращений:

ДИ – доверительный интервал

нРКИ – нерандомизированное исследование

РКИ – рандомизированное контролируемое исследование

Многие оригинальные исследования имеют сходные цели и задачи, но при этом проводятся разными коллективами, с разными пациентами, по иным протоколам и в разные временные промежутки. Результаты таких исследований могут быть разнообразны и противоречивы, что затрудняет принятие клинических решений. Благодаря концепции доказательной медицины разработаны инструменты, позволяющие

объединить результаты многочисленных исследований, в определенной мере различающихся между собой [1]. Проверая гипотезы в разных популяциях, мы получаем бесценные данные о том, что: а) эффект прослеживается у разных групп (или, наоборот, есть только в конкретных группах); б) существует вариативность эффекта. Большое количество сходных исследований, по сути, являются репликациями одного

большого эксперимента. Соответственно большее количество репликаций повышает мощность и степень доверия к результатам.

Существуют несколько основных инструментов, чтобы вынести суждение о совокупных результатах неких единообразных исследований.

- **Систематические обзоры.** Систематические обзоры включают все исследования, которые возможно найти, соответствующие неким строгим критериям включения. Данные критерии включения направлены на соблюдение ряда стандартов, предъявляемых к планированию исследований, их проведению и стандартным операционным процедурам и результатам (например, исследования по артериальной гипертензии, где артериальное давление оценивалось в соответствии с международными рекомендациями по кардиологии; исследования определенной манипуляции в хирургии, которая выполнялась строго в соответствии с четкими рекомендациями международных ассоциаций).
- **Мета-анализы.** Главным отличием мета-анализа от систематических обзоров является не только рассмотрение результатов ряда исследований, но и количественная оценка их результатов. В самом деле, нам необходимо знать не только то, что препарат/вмешательство обладает эффектом, но и оценить размер этого эффекта, и широту его варируемости.

В настоящей публикации мы затронем базовые аспекты проведения мета-анализов, что делается

после того, как проведен литературный поиск и фактически проведена большая часть работы для систематического обзора.

ВКЛЮЧЕНИЕ ИССЛЕДОВАНИЙ В МЕТА-АНАЛИЗ

Рекомендации PRISMA

Детальное описание критериев и процесса литературного поиска не является темой настоящей публикации, однако важно помнить, что любое создание систематического обзора состоит из ряда ступеней. Эти ступени объединены в стандартную схему, получившую название «диаграмма PRISMA» (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, предпочтительные элементы отчетности для систематических обзоров и мета-анализов), по названию соответствующих рекомендаций [2] (рис. 1).

Собственно, мета-анализ начинается после того, как завершено включение и доступны исследования, из которых будет выполнена экстракция данных.

Основными рекомендациями по предварительным и дальнейшим шагам, представляющим для читателя практический интерес, являются:

- Руководство Кокрановского общества по составлению систематических обзоров (Cochrane Handbook for Systematic Reviews of Interventions)¹ – данный документ размещен в интернете и представляет собой структурированное руководство с освещением практически всех

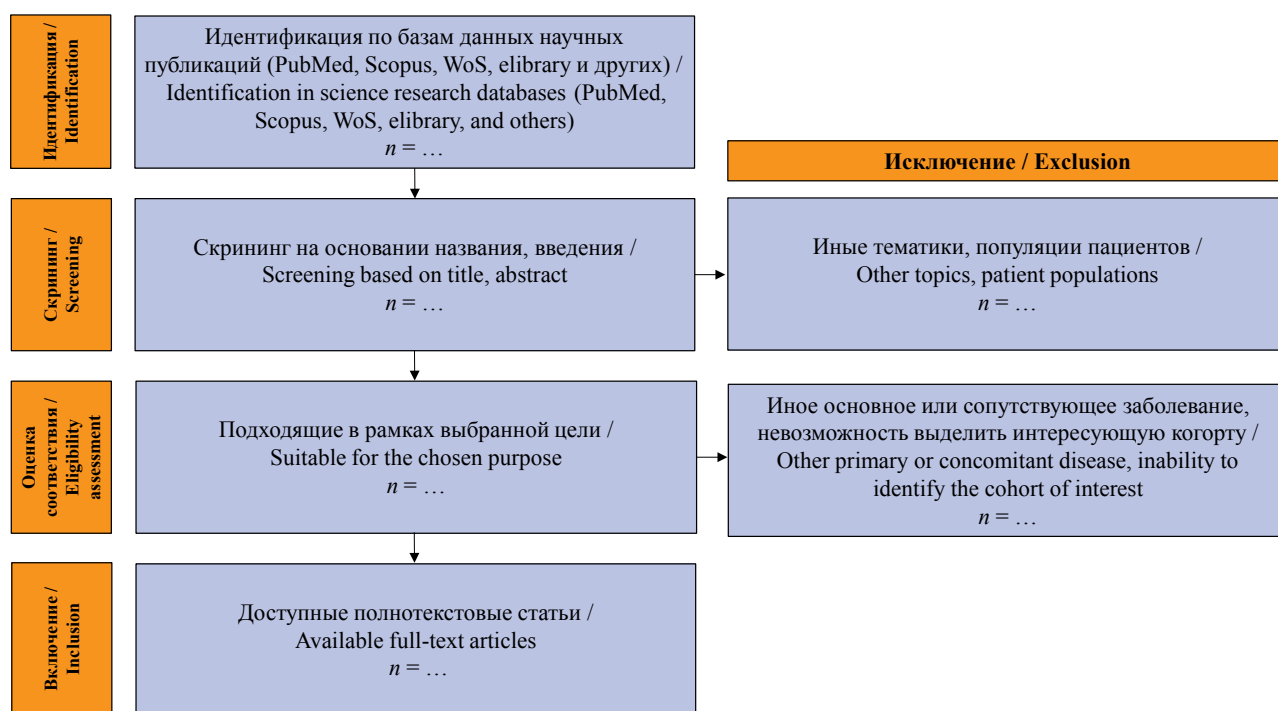


РИС. 1. Шаблон потоковой диаграммы PRISMA, адаптирована из M.J. Page и соавт. [2]

FIG. 1. PRISMA flow diagram template, adapted from M.J. Page et al. [2]

¹ Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from <https://training.cochrane.org/handbook/current>

аспектов литературного поиска, проведения мета-анализа, базовых навыков по математической статистике, необходимых для такой работы.

- Рекомендации PRISMA, о которых мы упоминали выше [2].

Представленные рекомендации помогут соблюсти высокие стандарты при написании систематических обзоров и уменьшить число возможных ошибок и неточностей, которые впоследствии могут затруднить продолжение работы. В связи с тем что мета-анализы широко используются фармацевтическими компаниями, а также иными коммерческими структурами, соблюдение указанных в рекомендациях стандартов проверяется рецензентами очень тщательно как при рассмотрении, так и при публикации статей. Строгое соблюдение регламента и рекомендаций является залогом успешной публикации.

ОЦЕНКА ВОЗМОЖНЫХ СМЕЩЕНИЙ В ПУБЛИКАЦИЯХ

После сбора всех подходящих публикаций для включения в систематический обзор или мета-анализ необходимо оценить их с точки зрения возможных смещений. К сожалению, источников потенциальных смещений может быть довольно много, поэтому были разработаны специальные инструменты, позволяющие исследователям проводить потенциальную оценку публикаций. Такие инструменты получили название *графики риска смещений*, или *risk of bias plots*.

Данные графики можно строить для исследований с различными типами дизайна, в первую очередь для рандомизированных контролируемых исследований (РКИ) (рис. 2) и для нерандомизированных исследований (неРКИ) [3, 4] (рис. 3). Графики, приведенные ниже, получили название «светофоры». Основная идея сводится к тому, что исследователи, проводящие мета-анализ с включением РКИ (рис. 2), рассматривают каждое исследование по отдельности и оценивают, насколько велики риски, ассоциированные с 5 доменами:

- рандомизацией;
- вмешательством;
- пропущенными данными;
- оценкой конечной точки;
- представлением результатов.

Риск оценивается как «высокий», «средний» и «низкий».

В случае если в мета-анализ включаются неРКИ, оцениваются риски, связанные с 7 доменами:

- конфаундингом (confounding, влияние вмешивающихся факторов);
- отбором и включением пациентов;
- вмешательством;
- расхождением с протоколом;
- пропущенными данными;
- оценкой конечной точки;
- представлением результатов.

Оценка риска интерпретируется как потенциально «низкая», «умеренная», «высокая» и «очень высокая».

Домены риска смещений / Risk of bias domains

	Д1 / D1	Д2 / D2	Д3 / D3	Д4 / D4	Д5 / D5	Общая оценка / Overall
Исследование 1 / Study 1	+	+	+	+	+	+
Исследование 2 / Study 2	X	X	+	+	-	+
Исследование 3 / Study 3	+	-	-	X	+	-
Исследование 4 / Study 4	+	+	X	+	-	X
Исследование 5 / Study 5	-	+	-	+	+	-

Оценка риска / Judgement

X высокий / high
 - средний / some concerns
 + низкий / low

РИС. 2. Диаграмма светофор для рандомизированных контролируемых исследований, адаптирована из L.A. McGuinness и соавт. [5]

FIG. 2. Traffic light plot for randomized controlled trials, adapted from L.A. McGuinness et al. [5]

Примечание: риск смещения, ассоциированный с доменами: Д1 – рандомизацией; Д2 – вмешательством; Д3 – пропущенными данными; Д4 – оценкой конечной точки; Д5 – представлением результатов.

Note: risk of bias associated with the domains: D1 – randomization; D2 – deviations from intended interventions; D3 – missing data; D4 – measurement of outcome; D5 – selection of reported results.

Домены риска смещений / Risk of bias domains

	Д1 / D1	Д2 / D2	Д3 / D3	Д4 / D4	Д5 / D5	Д6 / D6	Д7 / D7	Общая оценка / Overall
Исследование 1 / Study 1	!	+	!	!	+	+	X	!
Исследование 2 / Study 2	X	X	+	+	-	+	-	X
Исследование 3 / Study 3	+	+	X	!	-	+	!	+
Исследование 4 / Study 4	!	!	+	!	!	!	!	!
Исследование 5 / Study 5	!	-	+	+	+	+	!	-

Оценка риска / Judgement

! очень высокий / critical X высокий / serious - умеренный / moderate + низкий / low

РИС. 3. Диаграмма светофор для когортных исследований, адаптирована из L.A. McGuinness и соавт. [5]**FIG. 3.** Traffic light plot for cohort studies, adapted from L.A. McGuinness et al. [5]

Примечание: риск смещения, ассоциированный с доменами: Д1 – конфаундинг (влияние вмешивающихся факторов); Д2 – отбором и включением пациентов; Д3 – вмешательством; Д4 – расхождением с протоколом; Д5 – пропущенными данными; Д6 – оценкой конечной точки; Д7 – представлением результатов.

Note: risk of bias associated with the domains: D1 – confounding; D2 – selection of participants; D3 – classification of interventions; D4 – deviations from intended interventions; D5 – missing data; D6 – measurement of outcome; D7 – selection of reported results.

Данный инструмент позволяет критически подходить к полученным в мета-анализе результатам и рассматривать исследования с высоким риском смещения как менее надежные. Детальное описание возможностей данного инструмента приведено на специализированном сайте².

РАЗМЕР ЭФФЕКТА В МЕТА-АНАЛИЗЕ

Исследования, объединенные в мета-анализе, обычно оценивают результаты по идентичной конечной точке. Данная оценка есть эффект, который был достигнут, или *наблюдаемый эффект*. Эффект традиционно обозначается греческой буквой «тета» θ , наблюдаемый эффект по каждому из k исследований обычно обозначается, как θ_k . Понятие эффекта, его размера и его оценки описаны в нашей публикации о тестировании статистических гипотез [6].

Существуют две основные концепции, позволяющие описать и оценить эффект нескольких исследований. Обе эти концепции привязаны к соответствующим статистическим моделям с *фиксированными* и со *случайными* эффектами.

Модель с фиксированными эффектами

Эта модель подразумевает, что включенные в мета-анализ исследования очень похожи друг на друга в отношении дизайна, количества пациентов, методологии проведения, оценки результатов и других параметров.

Исследования представляются сходными настолько, что их результаты, или эффекты $\theta_{1,2,3,...,k}$, считают единой выборкой из одной генеральной совокупности всех возможных аналогичных исследований.

Вероятностное распределение такой совокупности имеет математическое ожидание (среднее взвешенное по вероятностям возможных значений), представляющее собой некий *истинный размер эффекта* $\hat{\theta}$. Соответственно каждое исследование представляет собой элемент совокупности, а несколько случайно взятых из такого распределения исследований (мета-анализ) представляют собой банальную выборку.

Наблюдаемый эффект в каждом исследовании k будет отличаться от истинного на величину ошибки:

$$\hat{\theta} = \theta_k + \epsilon_k.$$

Соответственно мы считаем, что среди нескольких исследований наиболее точными являются те, где величина выборочной ошибки ϵ наименьшая (рис. 4).

Выборка из нескольких исследований $\{1,2,3,...,k\}$ должна иметь некую центральную тенденцию или математическое ожидание, отражающее истинный размер эффекта. Таким образом, используя определение математического ожидания как среднее взвешенное, получаем:

$$\hat{\theta} = \frac{w_1\theta_1 + w_2\theta_2 + w_3\theta_3 + \dots + w_k\theta_k}{w_1 + w_2 + w_3 + \dots + w_k},$$

² <https://www.riskofbias.info/> (access date 01.11.2022).

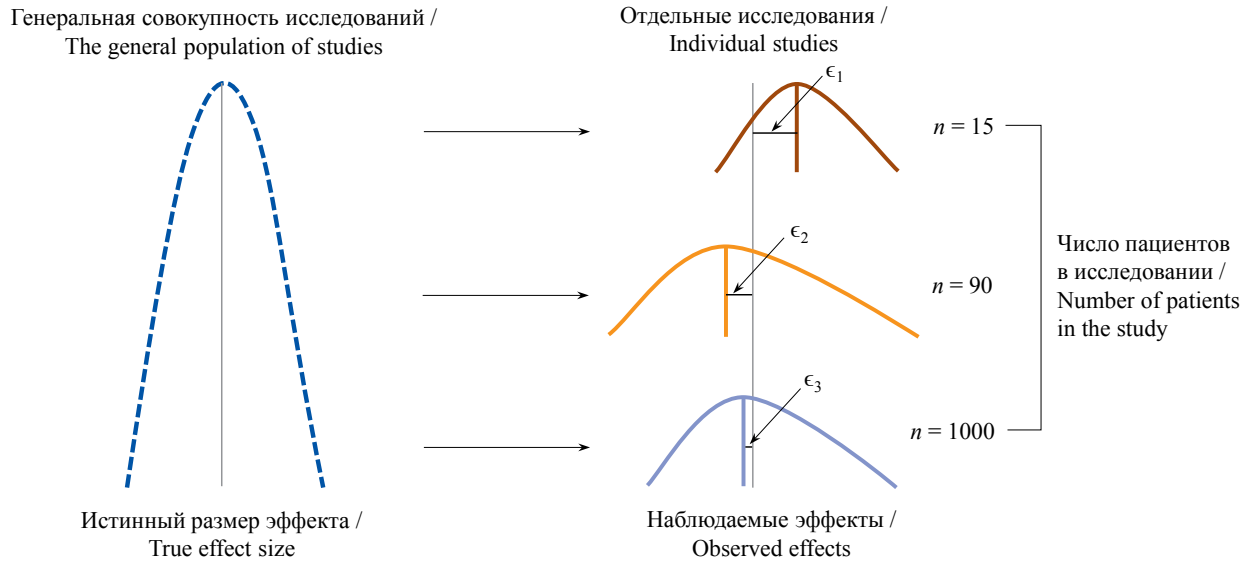


РИС. 4. Схематичное представление истинного и наблюдаемых эффектов в исследованиях при использовании модели с фиксированными эффектами

FIG. 4. Schematic representation of true and observed effects in studies using a fixed effects model

Примечание: ϵ_k – величина выборочной ошибки; серая линия – истинный размер эффекта; цветные прямые линии – наблюдаемые эффекты.
Note: ϵ_k – sampling error; grey line – true effect size; colored straight lines – observed effects.

где

- $\hat{\theta}$ – взвешенный размер эффекта по k исследованиям, являющийся результатом мета-анализа;
- θ_k – наблюдаемый эффект исследования k ;
- w_k – вес исследования k .

Из данного уравнения неясным остается, как получить вес каждого исследования. Как мы знаем, наблюдаемый эффект, полученный в исследовании k , является точечной оценкой (point estimate). В исследование k включено определенное количество пациентов, n . Мерой вариабельности эффекта θ_k будет его стандартная ошибка, рассчитанная как:

$$s_k = \frac{\sigma_k}{\sqrt{n}},$$

где

- σ – стандартное отклонение эффекта θ_k в исследовании k ;
- n – количество пациентов в исследовании k .

В модели с фиксированными эффектами один из способов расчета весов выполняется с помощью так называемого метода обратной дисперсии (inverse variance method):

$$w_k = \frac{1}{s_k^2},$$

где

- s_k^2 – квадрат стандартной ошибки эффекта θ_k в исследовании k ;
- w_k – вес исследования k .

Таким образом, в модели с фиксированными эффектами вес любого исследования обратно связан с величиной обратной ошибки эффекта в исследовании и напрямую – с количеством включенных в исследование пациентов. Допущением модели является

тот факт, что ничто, кроме количества пациентов, не может влиять на вес исследования.

Данная концепция, с одной стороны, представляется излишне упрощенной. В реальном мире исследования так или иначе отличаются друг от друга и данные различия обусловлены огромным количеством факторов помимо размера выборки.

Модель со случайными эффектами

Если в мета-анализе объединены исследования различного дизайна (РКИ, когортные и т.п.), проведенные в разные годы, в разных странах, в центрах с разными стандартами оказания медицинской помощи, с протоколами вмешательств по разным клиническим рекомендациям, требуется модель, которая будет учитывать различия не только по объему выборки, но и по вышеупомянутым факторам.

Такая модель предполагает, что эффект по каждому исследованию, включенному в мета-анализ, является выборкой из собственной совокупности размеров эффекта и отличается от математического ожидания собственной совокупности на величину ϵ_k (рис. 5).

Если мы включили в мета-анализ k исследований, следовательно, это k выборок из k различных генеральных совокупностей. При этом k совокупностей имеют свое распределение с математическим ожиданием в виде взвешенного размера эффекта, при этом каждая генеральная совокупность отличается от математического ожидания своего распределения на величину ξ_k . Такое общее распределение имеет точечную оценку, которая соответствует взвешенному эффекту в мета-анализе, $\hat{\theta}$, и дисперсию τ^2 .

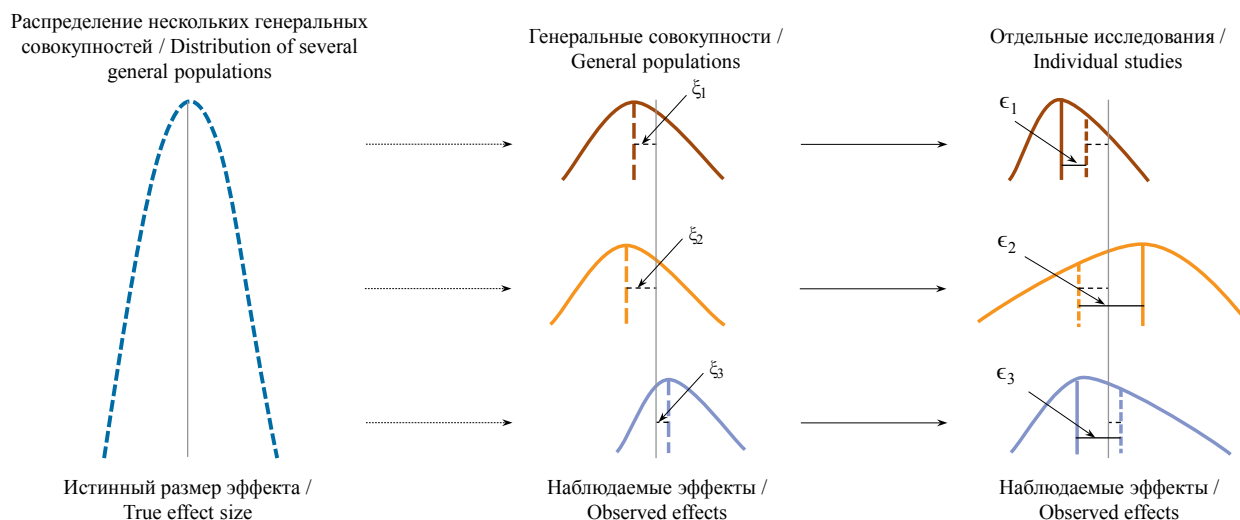


РИС. 5. Схематичное представление истинного и наблюдаемых эффектов в исследованиях при использовании модели со случайными эффектами

FIG. 5. Schematic representation of true and observed effects in studies using a random effects model

Примечание: ξ_k – величина ошибки каждой генеральной совокупности; ϵ_k – величина выборочной ошибки; серая линия – истинный размер эффекта; цветные пунктирные линии – наблюдаемые эффекты в генеральной совокупности; цветные прямые линии – в отдельных исследованиях.

Note: ξ_k – error of each general population; ϵ_k – sampling error; grey line – true effect size; colored dotted lines – observed effects in general populations; colored straight lines – in individual studies.

Таким образом, точечная оценка эффекта в каждом исследовании среди k отличается от взвешенного эффекта следующим образом:

$$\hat{\theta} = \theta_k + \epsilon_k + \xi_k.$$

Показатель ξ_k объединяет различия, которые не связаны с ошибкой выборки.

Модель, которая использует данную логику, получила название *модель со случайными эффектами*. Веса индивидуальных исследований в такой модели рассчитываются следующим образом:

$$w_k = \frac{1}{s_k^2 + \tau^2},$$

где

- s_k^2 – квадрат стандартной ошибки эффекта θ_k в исследовании k ;
- w_k – вес исследования k ;
- τ^2 – дисперсия распределения нескольких генеральных совокупностей.

В свою очередь, взвешенный эффект рассчитывается абсолютно так же, как и для модели с фиксированными эффектами:

$$\hat{\theta} = \frac{w_1\theta_1 + w_2\theta_2 + w_3\theta_3 + \dots + w_k\theta_k}{w_1 + w_2 + w_3 + \dots + w_k},$$

где

- $\hat{\theta}$ – взвешенный размер эффекта по k исследованиям, являющийся результатом мета-анализа;
- θ_k – наблюдаемый эффект исследования k ;
- w_k – вес исследования k .

Для вычисления показателя τ^2 используются различные математические подходы, наиболее

часто – DerSimonian-Laird, ограниченного максимального правдоподобия (Restricted Maximum Likelihood), максимального правдоподобия (Maximum Likelihood), процедура Paule-Mandel и другие [7–9]. Выбор конкретного метода зависит от типа оценки конечной точки и от конкретной ситуации, поэтому требует консультации с биостатистиком.

ОЦЕНКА ГЕТЕРОГЕННОСТИ

Мы выяснили, что включенные в мета-анализ исследования могут значительно различаться, более того, в зависимости от этих различий избирается та или иная модель анализа. Существует ли некая мера, с помощью которой можно оценивать степень различий? Можем ли мы как-то объяснить степень различий, и нужно ли это делать? Как определить, какие исследования больше отличаются от других? На все эти вопросы может дать ответ концепция о *гетерогенности*.

Гетерогенность зависит от многих причин, наиболее частые из которых:

- неверный подбор исследований для мета-анализа;
- наличие явных и скрытых *факторов-модераторов*, влияющих на взвешенный эффект (модератор фактически создает подгруппы с разным размером эффекта);
- небольшое количество исследований, используемых для мета-анализа.

Когда исследователи сталкиваются с избыточной гетерогенностью, например в ситуации, где у ряда исследований противоположное направление эффекта или когда крайне разный размер эффекта, необходимо разобраться, не имеет ли место ошибочное

включение исследований в мета-анализ. Попытка объединить исследования, в которых оценивались совершенно разные параметры, приведет к огромной гетерогенности и не ответит на вопрос мета-анализа. Результаты такого мета-анализа будут крайне сомнительны.

Если мы считаем, что на данном этапе ошибки нет, необходимо далее искать причину высокой гетерогенности. Например, при использовании нестандартизованных показателей необходимо попытаться использовать стандартизованные. Ряд показателей может быть измерен в разных шкалах или значительно различаться на включении между исследованиями – стандартизация позволяет сгладить такие различия.

Следующей причиной высокой гетерогенности является наличие скрытых или явных модераторов или конфаундеров. Например, при оценке распространенности сердечно-сосудистых заболеваний явными модераторами являются пол и возрастная группа пациентов. Включение модератора и оценка его влияния на эффект и гетерогенность проводится с помощью *метарегрессионного анализа, или метарегрессии*. Дальнейший анализ подгрупп позволяет значительно снизить гетерогенность.

Наконец, малое количество исследований в мета-анализе может привести к высокой гетерогенности.

По G. Rücker, существует 2 основных типа гетерогенности [10]:

- Гетерогенность, обусловленная дизайном или базовыми характеристиками. Причиной является попытка объединить в мета-анализе различные по дизайну исследования (включая тип исследования, характер вмешательства, способ оценки результатов, срок их оценки и другие параметры), а также исследования, сильно гетерогенные по когортам пациентов. Данный тип гетерогенности может приводить к *статистической гетерогенности*.
- Статистическая гетерогенность. Обусловлена точностью оценки и вариабельностью размера эффекта. Данную гетерогенность уже можно оценить количественно. Одной из причин (но совсем не всегда) статистической гетерогенности может быть гетерогенность, обусловленная дизайном.

Способы оценки гетерогенности

Cochran's Q

Мы рассмотрели оба типа моделей и поняли, что существует наблюдаемый эффект $\hat{\theta}_k$ некоего исследования k , а также взвешенный эффект, который мы рассчитываем $\hat{\theta}$ для всех исследований мета-анализа. Мы также помним, что у каждого исследования есть свой вес w_k . Отклонение наблюдаемого эффекта от взвешенного может быть направлено в любую сторону (может быть с любым знаком). Если мы возведем отклонение в квадрат, оно больше не будет связано к направлению.

Сумма взвешенных квадратов таких отклонений получила название Cochran's Q:

$$Q = w_1(\hat{\theta}_1 - \hat{\theta})^2 + w_2(\hat{\theta}_2 - \hat{\theta})^2 + w_3(\hat{\theta}_3 - \hat{\theta})^2 + \dots + w_n(\hat{\theta}_n - \hat{\theta})^2.$$

Мы можем рассчитать отклонение наблюдаемого эффекта от взвешенного по всем исследованиям.

Распределение Cochran's Q сильно приближено к распределению χ^2 с $K - 1$ степенями свободы, где K – количество исследований в мета-анализе.

Cochran's Q будет расти при увеличении количества исследований в мета-анализе, а также при наличии в нем крупных исследований с большим количеством пациентов.

I²-статистика Хиггинса и Томпсона

Данная статистика рассчитывается с помощью Cochran's Q и указывает на % гетерогенности, обусловленный ошибкой ξ_k (ошибкой, не связанной с размером выборки).

Нулевая гипотеза состоит в том, что гетерогенность отсутствует и Cochran's Q следует распределению χ^2 с $K - 1$ степенями свободы (ожидаемая гетерогенность). Но у нас имеется и наблюдаемая гетерогенность Q. Тогда отклонение наблюдаемой гетерогенности от ожидаемой составляет:

$$I^2 = \frac{Q - (K - 1)}{Q}$$

и выражается в долях единицы или в процентах.

Принято выделять низкую (до 25%), умеренную (25–75%) и высокую (75% и выше) гетерогенность [11].

I²-статистика

Данная статистика рассчитывает отношение Q-статистики к $K - 1$: при отсутствии гетерогенности значение стремится к 1. Более высокие значения говорят о наличии гетерогенности между исследованиями.

$$I^2 = \frac{Q}{K - 1}.$$

Дисперсия гетерогенности τ^2

Истинный взвешенный эффект имеет собственную дисперсию τ^2 и стандартное отклонение τ . Данная статистика хорошо подходит для оценки меры гетерогенности, имея ту же размерность, что и эффект в исследованиях в мета-анализе. Зная рассчитанный взвешенный размер эффекта $\hat{\theta}$ в мета-анализе, мы можем оценить 95% доверительный интервал (ДИ) истинного эффекта как $\hat{\theta} \pm 1,96 \times \tau$.

АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ

Анализ чувствительности показывает, какое влияние могут оказать отдельные исследования на взвешенный эффект, или, другими словами, насколько результаты мета-анализа устойчивы.

Одним из наиболее часто используемых способов оценки чувствительности является метод leave-one-out.

Каждое исследование по одному удаляется из мета-анализа, после чего размер взвешенного эффекта и гетерогенность пересчитываются. Серьезные изменения по величине размера эффекта и снижение гетерогенности указывают на то, что именно исключенное исследование оказывает существенное влияние на общий результат. Если первоначально такое исследование в графике-«светофоре» было оценено исследователями как обладающее высоким/умеренным риском смещения, оно может быть исследованием-выбросом, и необходимо рассматривать целесообразность его присутствия в мета-анализе.

Графически результаты анализа чувствительности представлены на рисунке 6.

В примере на рисунке 6 взвешенный эффект, полученный в мета-анализе, является разницей средних (mean difference) и составил 0,16 [0,1; 0,23], мы видим, что при исключении поочередно исследований из мета-анализа взвешенный эффект существенно не меняется. При этом исключение исследования Protocol 162A существенно снижает гетерогенность, по статистике I^2 до 5%. Данное исследование требует пристального внимания, так как его присутствие обуславливает высокие значения гетерогенности во всем мета-анализе.

В общем случае анализ чувствительности направлен на оценку того, как меняется взвешенный эффект при исключении исследований, которым при оценке потенциальных рисков смещения был присвоен «высокий» и «очень высокий» уровень риска. При возникновении существенного влияния на эффект при исключении таких исследований (например, в ситуации, когда после исключения исследований с высоким риском 95% ДИ нового взвешенного эффекта перестает включать точечную оценку взвешенного эффекта до исключения исследований)

необходимо пересмотреть необходимость включения исследований с высоким риском в мета-анализ.

ФОРЕСТ-ДИАГРАММА (FOREST PLOT)

Наиболее общим способом суммировать результаты мета-анализа является форест-диаграмма (forest plot) – специфический график, отображающий исследования, вошедшие в анализ, эффект в каждом из них, взвешенный эффект, а также ряд дополнительных характеристик, например, веса каждого исследования, показатели гетерогенности, тип выбранной статистической модели (фиксированные или случайные эффекты). Неотъемлемой частью forest plot является диаграмма, дублирующая размеры эффекта и их 95% ДИ, а также взвешенный эффект и его 95% ДИ.

Довольно часто размер точек на графике, характеризующих определенное исследование, связан с весом данного исследования (соответственно, точки наибольшего размера связаны с исследованиями с наибольшим весом).

Типичный forest plot показан на рисунке 7. Данные исследования взяты из материалов, сопровождающих библиотеку meta языка программирования R [12].

На рисунке 7 мы видим названия исследований и год публикации результатов, характеристики основных и контрольных групп (количества, средние и стандартные отклонения – необходимые нам данные для расчета стандартной ошибки), размер эффекта (в данном случае разница средних, mean difference, MD) и его 95% ДИ, веса в моделях с фиксированными и случайными эффектами, взвешенный эффект для обоих типов моделей, а также показатели гетерогенности.

Также мы видим диаграмму с отрисовкой все тех же размеров эффекта в исследованиях (в виде

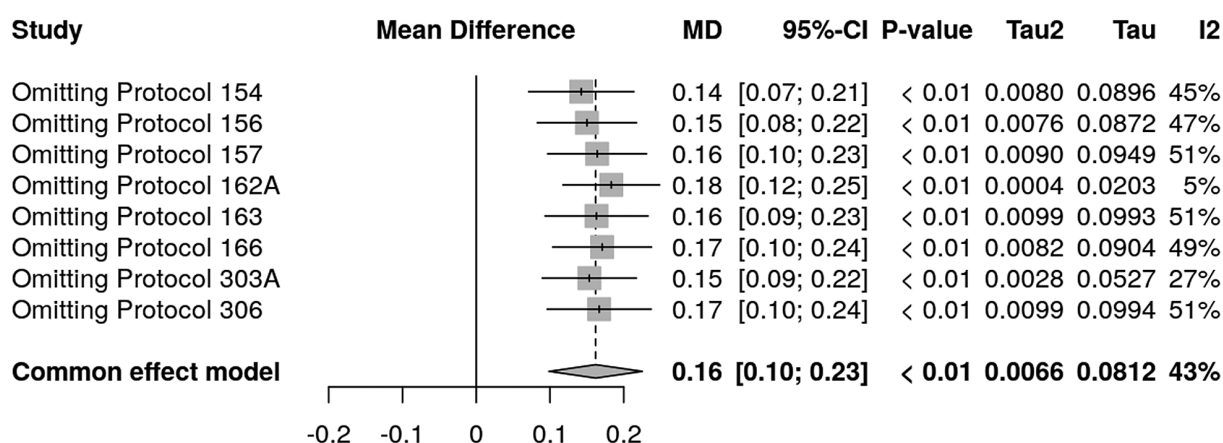


РИС. 6. Форест-диаграмма, демонстрирующая анализ чувствительности, адаптирована из S. Balduzzi и соавт. [12]

FIG. 6. Forest plot showing sensitivity analysis, adapted from S. Balduzzi et al. [12]

Примечание: study – исследование; MD – Mean difference, разница средних; CI – confidence interval, доверительный интервал; P-value – значение p; I2 – Higgin's & Thompson's I^2 statistic / I^2 статистика Хиггинса и Томпсона / I^2 -статистика; common (fixed) effect model – модель с фиксированными эффектами.

Note: MD – Mean difference; CI – confidence interval; I2 – Higgin's & Thompson's I^2 statistic / I^2 statistic.

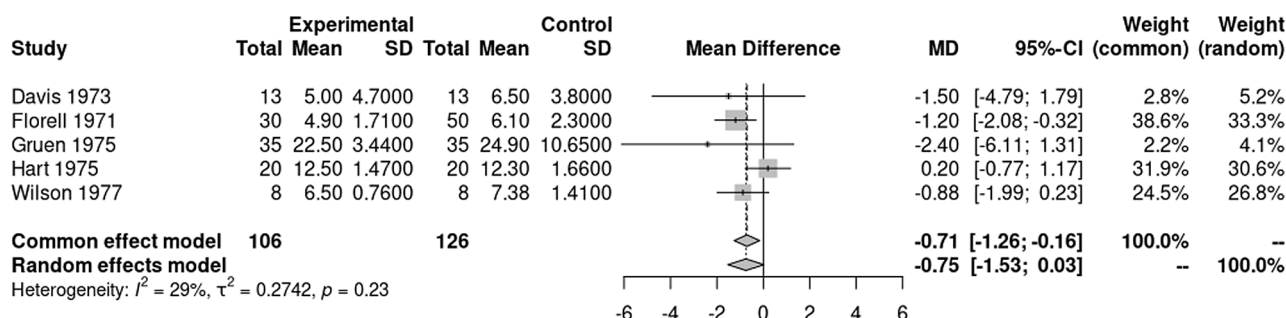


РИС. 7. Forest-диаграмма, демонстрирующая взвешенный эффект, адаптирована из S. Balduzzi и соавт. [12]

FIG. 7. Forest plot showing the weighted effect, adapted from S. Balduzzi et al. [12]

Примечание: study – исследование; experimental – экспериментальная группа; control – контрольная группа; total – общее значение; mean – среднее значение; SD – standard deviation, стандартное отклонение; MD – mean difference, разница средних; CI – confidence interval, доверительный интервал; weight (common) – веса в модели с фиксированными эффектами; weight (random) – веса в модели со случайными эффектами; common (fixed) effect model – модель с фиксированными эффектами; random effect model – модель со случайными эффектами; heterogeneity – гетерогенность; I^2 – Higgin's & Thompson's I^2 statistic / I^2 statistic, I^2 -статистика Хиггинса и Томпсона / I^2 -статистика; τ^2 – Tau-squared, Tay-квадрат; P-value – значение p .

Note: SD – standard deviation; MD – mean difference; CI – confidence interval; I^2 – Higgin's & Thompson's I^2 statistic / I^2 statistic; τ^2 – Tau-squared.

квадратов, размер которых указывает на веса) и взвешенные эффекты для моделей с фиксированными и случайными эффектами (в виде ромбов).

Вертикальная сплошная черта обозначает размер эффекта, равный нулю. Так как мы имеем дело с разницей средних, если 95% ДИ отдельных исследований или взвешенных эффектов включает ноль, значит высока вероятность, что наблюдаемые точечные оценки случайны.

В настоящем примере единственное исследование, где наблюдается значимый эффект, – Florell 1971, его вес в модели как с фиксированными, так и со случайными эффектами наивысший (38,6 и 33,3% соответственно). По результатам мета-анализа модель с фиксированными эффектами показывает наличие значимого эффекта, разница средних составила -0,71 [-1,26; -0,16], при этом результаты модели со случайными эффектами оказались незначимыми: точечная оценка составила -0,75, а 95% ДИ включает ноль [-1,53; 0,03].

ВКЛАД АВТОРОВ

А.Ю. Суворов, И.В. Латушкина и К.А. Гуляева в равной степени внесли вклад в эту работу и должны считаться первыми соавторами. А.Ю. Суворов, И.В. Латушкина, К.А. Гуляева, Н.М. Буланов, М.Ю. Надинская и А.А. Заикин участвовали в написании текста рукописи. А.Ю. Суворов, И.В. Латушкина, Н.М. Буланов и А.А. Заикин выполняли поиск и анализ литературы по теме обзора. А.Ю. Суворов разработал общую концепцию статьи, А.А. Заикин осуществлял руководство ее написанием. Все авторы участвовали в обсуждении и редактировании работы. Все авторы утвердили окончательную версию публикации.

ЗАКЛЮЧЕНИЕ

В настоящей части статьи мы познакомились с этапами включения исследований в мета-анализ, рассмотрели существующие рекомендации, с которыми необходимо познакомиться при написании мета-анализа, детально разобрали процесс создания весов, различные типы моделей, используемых в мета-анализах. Также мы познакомились с понятием гетерогенности и способом ее вычисления. Были представлены основные графики, публикуемые в мета-анализах, такие как forest plot и график leave-one-out для анализа чувствительности.

В следующей публикации мы рассмотрим, как происходит анализ подгрупп и концепцию метарегрессии, а также научимся оценивать публикационное смещение, визуально и математическими методами. Дополнительно мы вспомним, как наиболее часто оценивается эффект в мета-анализе, и остановимся на стандартизованных показателях и оценке для них стандартной ошибки.

AUTHOR CONTRIBUTIONS

Alexander Yu. Suvorov, Irina V. Latushkina and Kseniya A. Gulyaeva equally contributed to this work and should be considered the first co-authors. Alexander Yu. Suvorov, Irina V. Latushkina, Kseniya A. Gulyaeva, Nikolay M. Bulanov, Maria Yu. Nadinskaia and Alexey A. Zaikin participated in writing the text of the manuscript. Alexander Yu. Suvorov, Irina V. Latushkina, Nikolay M. Bulanov and Alexey A. Zaikin searched and analyzed the literature on the topic of the review. Alexander Yu. Suvorov developed the general concept of the article, Alexey A. Zaikin supervised writing the text of the manuscript. All authors participated in the discussion and editing of the work. All authors approved the final version of the publication.

ЛИТЕРАТУРА / REFERENCES

- 1 Sackett D.L., Rosenberg W.M., Gray J.A., et al. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996 Jan 13; 312(7023): 71–72. <https://doi.org/10.1136/bmj.312.7023.71>. PMID: 8555924
- 2 Page M.J., McKenzie J.E., Bossuyt P.M., et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews *BMJ* 2021; 372: n71. <https://doi.org/10.1136/bmj.n71>
- 3 Sterne J.A.C., Savović J., Page M.J., et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019 Aug 28; 366: 14898. <https://doi.org/10.1136/bmj.14898>. PMID: 31462531
- 4 Sterne J.A., Hernán M.A., Reeves B.C., et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016 Oct 12; 355: i4919. <https://doi.org/10.1136/bmj.i4919>. PMID: 27733354
- 5 McGuinness L.A., Higgins J.P.T. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods*. 2021 Jan; 12(1): 55–61. <https://doi.org/10.1002/jrsm.1411>. Epub 2020 May 6. PMID: 32336025
- 6 Suvorov A.Yu., Bulanov N.M., Shvedova A.N., et al. Statistical hypothesis testing: general approach in medical research. *Sechenov Medical Journal*. 2022; 13(1): 4–13. <https://doi.org/10.47093/2218-7332.2022.426.08>
- 7 Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, 2005; 30(3): 261–293. <https://doi.org/10.3102/10769986030003261>
- 8 Veroniki A.A., Jackson D., Viechtbauer W., et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods*. 2016 Mar; 7(1): 55–79. <https://doi.org/10.1002/jrsm.1164>. Epub 2015 Sep 2. PMID: 26332144
- 9 Langan D., Higgins J.P.T., Jackson D., et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods*. 2019 Mar; 10(1): 83–98. <https://doi.org/10.1002/jrsm.1316>. Epub 2018 Sep 6. PMID: 30067315
- 10 Rücker G., Schwarzer G., Carpenter J.R., Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol*. 2008 Nov 27; 8: 79. <https://doi.org/10.1186/1471-2288-8-79>. PMID: 19036172
- 11 Higgins J.P., Thompson S.G. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002 Jun 15; 21(11): 1539–1558. <https://doi.org/10.1002/sim.1186>. PMID: 12111919
- 12 Balduzzi S., Rücker G., Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health*. 2019 Nov; 22(4): 153–160. <https://doi.org/10.1136/ebmental-2019-300117>. Epub 2019 Sep 28. PMID: 31563865

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Суворов Александр Юрьевич, канд. мед. наук, главный статистик Центра анализа сложных систем ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0002-2224-0019>

Латушкина Ирина Викторовна✉, младший научный сотрудник Центра анализа сложных систем ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0001-8885-6062>

Гуляева Ксения Александровна, аспирант кафедры пропедевтики внутренних болезней, гастроэнтерологии и гепатологии ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0002-3462-0123>

Буланов Николай Михайлович, канд. мед. наук, доцент кафедры внутренних, профессиональных болезней и ревматологии ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0002-3989-2590>

Надинская Мария Юрьевна, канд. мед. наук, доцент кафедры пропедевтики внутренних болезней, гастроэнтерологии и гепатологии ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0002-1210-2528>

Заикин Алексей Анатольевич, канд. физ.-мат. наук, заместитель директора Центра анализа сложных систем ФГАОУ ВО «Первый МГМУ им. И.М. Сеченова» Минздрава России (Сеченовский Университет).

ORCID: <https://orcid.org/0000-0001-7540-1130>

Alexander Yu. Suvorov, Cand. of Sci. (Medicine), Chief Statistician, Centre for Analysis of Complex Systems, Sechenov First Moscow State Medical University (Sechenov University).

ORCID: <https://orcid.org/0000-0002-2224-0019>

Irina V. Latushkina✉, junior researcher, Centre for Analysis of Complex Systems, Sechenov First Moscow State Medical University (Sechenov University)

ORCID: <https://orcid.org/0000-0001-8885-6062>

Kseniya A. Gulyaeva, postgraduate student, Department of Propaedeutics of Internal Diseases, Gastroenterology and Hepatology, Sechenov First Moscow State Medical University (Sechenov University)

ORCID: <https://orcid.org/0000-0002-3462-0123>

Nikolay M. Bulanov, Cand. of Sci. (Medicine), Associate Professor, Department of Internal, Occupational Diseases and Rheumatology, Sechenov First Moscow State Medical University (Sechenov University).

ORCID: <https://orcid.org/0000-0002-3989-2590>

Maria Yu. Nadinskaia, Cand. of Sci. (Medicine), Associate Professor, Department of Propaedeutics of Internal Diseases, Gastroenterology and Hepatology, Sechenov First Moscow State Medical University (Sechenov University).

ORCID: <https://orcid.org/0000-0002-1210-2528>

Alexey A. Zaikin, Cand. of Sci. (Physics and Mathematics), Deputy Director, Centre for Analysis of Complex Systems, Sechenov First Moscow State Medical University (Sechenov University).

ORCID: <https://orcid.org/0000-0001-7540-1130>

✉ Автор, ответственный за переписку / Corresponding author