

Оценка риска систематических ошибок в одномоментных исследованиях диагностических тестов: русскоязычная версия вопросника QUADAS

О. Ю. Реброва¹, В. К. Федяева^{2, 3}

¹ Российский национальный исследовательский медицинский университет им. Н. И. Пирогова, Москва, Россия

² Центр экспертизы и контроля качества медицинской помощи Министерства здравоохранения РФ, Москва, Россия

³ Научно-исследовательский финансовый институт Министерства финансов РФ, Москва, Россия

В статье приведена русскоязычная версия вопросника, предназначенного для оценки риска систематических смещений в одномоментных исследованиях диагностических тестов. Данная версия является адаптированным переводом международного консенсусного вопросника QUADAS. Риск систематических смещений в оригинальной шкале не стратифицирован, однако мы предлагаем оценивать общий риск систематических смещений как низкий, средний или высокий в зависимости от суммы баллов в соответствии с критериями вопросника. При общей оценке методологического качества исследования следует учитывать не только риск систематических ошибок, но и риск некорректности выполненного статистического анализа.

КЛЮЧЕВЫЕ СЛОВА: одномоментное исследование, диагностический тест, диагностическая точность, референсный тест, методологическое качество, вопросник, систематическая ошибка, оценка, QUADAS.

Assessment of Risk of Bias in the Cross-Sectional Studies of Diagnostic Tests: the Russian-Language Version of the Questionnaire QUADAS

O. Yu. Rebrova¹, V. K. Fediaeva^{2, 3}

¹ N. I. Pirogov Russian National Research Medical University, Moscow, Russia

² Center of Expertise and Quality Control of Medical Care, Ministry of Health of the Russian Federation, Moscow, Russia

³ Financial Scientific Research Institute at the Ministry of Finances of Russian Federation, Moscow, Russia

The article presents the Russian-language version of the questionnaire to assess the risk of systematic bias in cross-sectional studies of diagnostic tests. This version is the adapted translation of the international consensus questionnaire QUADAS. The risk of systematic bias in the original scale is not stratified, however, we propose to evaluate the overall risk of systematic bias as low, medium or high, depending on the score sum according to the criteria of the questionnaire. To assess the overall methodological quality not only the risk of systematic bias should be assessed, but also the risk of incorrect statistical analysis.

KEYWORDS: cross-sectional study, diagnostic test, diagnostic accuracy, reference test, systematic bias, assessment, QUADAS.

Исследования диагностических тестов в большинстве случаев имеют целью оценку диагностической точности теста относительно «золотого стандарта диагностики» или другого референсного теста. Показателями точности являются:

- устойчивые операционные характеристики теста (не зависящие от преваленса, т. е. распространен-

ности диагностируемого состояния в целевой популяции): чувствительность, специфичность;

- зависящие от преваленса характеристики теста: прогностическая ценность положительного и отрицательного результата.

Такие исследования в идеале должны иметь дизайн одномоментного (cross-sectional) исследования, т. к. все

Таблица 1. Шкала для оценки риска систематических ошибок в одномоментном исследовании диагностического теста, разработанная нами на основе вопросника QUADAS

Ответы	Вопросы и пояснения	Баллы
1. Была ли выборка пациентов репрезентативной по отношению к целевой популяции?*		
Да	Выборка пациентов соответствовала заранее установленным критериям включения/исключения Набор пациентов был сплошным либо включались случайные выборки из последовательных серий	1
Нет	Выборка нерепрезентативна по отношению к целевой популяции либо группы (с целевым состоянием и без него) были набраны из разных популяций	0
Неясно	Недостаточно информации о выборке пациентов или методе формирования выборки	0
2. Достаточно ли хорошо референсный тест классифицирует целевые состояния?*		
Да	Референсный тест соответствует заранее установленным критериям	1
Нет	Референсный тест не соответствует заранее установленным критериям либо такие критерии не установлены	0
Неясно	Неясно, какой именно референсный тест использован	0
3. Был ли промежуток времени между выполнением референсного и изучаемого тестов достаточно коротким, чтобы быть уверенным, что изучаемое состояние не изменилось за это время?*		
Да	Интервал времени был приемлемым по крайней мере для достаточно высокой доли пациентов	1
Нет	Интервал времени был больше приемлемого у большой доли пациентов	0
Неясно		0
4. Выполнена ли верификация референсным тестом для всей выборки или ее случайной части?		
Да	Все пациенты либо случайно отобранные пациенты обследованы референсным тестом	1
Нет	Некоторые пациенты не обследованы референсным тестом, и их отбор не был случайным	0
Неясно		0
5. Обследованы ли пациенты одним и тем же референсным тестом независимо от результатов изучаемого теста?		
Да		1
Нет		0
Неясно		0
6. Применялся ли референсный тест независимо от исследуемого (исследуемый тест не являлся компонентом референсного)?		
Да		1
Нет		0
Неясно		0
7. Были ли результаты референсного теста интерпретированы без информации о результате исследуемого теста?		
Да	Да, интерпретированы вслепую, и это обеспечивалось либо последовательностью выполнения тестов, либо протоколом исследования	1
Нет		0
Неясно		0
8. Были ли результаты исследуемого теста интерпретированы без знания результатов референсного теста?		
Да	Да, интерпретированы вслепую, и это обеспечивалось либо последовательностью выполнения тестов, либо протоколом исследования	1
Нет		0
Неясно		0
9. Была ли при интерпретации результатов исследуемого и референсного тестов доступна вся та клиническая информация, которая должна быть доступна при практическом применении теста?*		
Да	Да, в соответствии с заранее установленными критериями	1
Нет	Нет, клиническая информация была изъята или, наоборот, дополнительно предоставлена	0
Неясно		0
10. Указано ли число неинтерпретируемых/промежуточных результатов теста?		
Да	Да, указано, либо таких результатов не было (число интерпретированных тестов равно числу пациентов)	1
Нет	Нет (указывается, что такие результаты были, но не приведено их число)	0
Неясно		0
11. Объяснены ли выбывания (случаи, когда не применен один из тестов) из исследования?		
Да	Да, есть потоковая диаграмма с указанием причин выбывания или исключения, или выбываний не было	1
Нет	Нет (выбывание было, но не объяснено)	0
Неясно		0
Максимальная сумма баллов		11

* Для вопросов 1, 2, 3, 9 рекомендуется разработать правила оценки в рамках конкретной проблемы. В дальнейшем они должны быть включены в публикацию.

Таблица 2. Дополнительные аспекты, учитываемые при оценке риска систематических ошибок в исследовании диагностического теста

Номер	Вопрос
1	Были ли точки разделения дифференцируемых состояний (граничные значения количественного теста) установлены до начала исследования?
2	Соблюдались ли одна и та же технология выполнения исследуемого теста на протяжении исследования?
3	Дано ли четкое определение позитивного результата?
4	Были ли исполнители достаточно обучены выполнению теста?
5	Сохранялась ли неизменность лечения на протяжении исследования?
6	Приведены ли сведения о вариабельности наблюдения и находились ли они в приемлемом диапазоне?
7	Приведены ли сведения об инструментальной вариабельности и находились ли они в приемлемом диапазоне?
8	Были ли цели исследования определены заранее?
9	Установлено ли, что исследование не имело финансирования со стороны бизнеса и что у исследователей не было конфликта интересов?

другие дизайны чреватые более высоким риском систематических смещений результатов. При этом в исследованиях диагностических тестов крайне важна не только внутренняя валидность исследования (определяемая риском систематических ошибок и вероятностью некорректности результатов статистического анализа), но и внешняя валидность. Последняя определяется прежде всего исследуемой популяцией, поскольку от ее характеристик (преваленса диагностируемых состояний и проч.) существенно зависит прогностическая ценность теста. В ситуации высокого преваленса (обычно в клинической ситуации) диагностические тесты разрабатываются собственно как диагностические, при этом рекомендуется оптимизировать (максимизировать) диагностическую чувствительность и прогностическую ценность положительного результата. Если же тест применяется в условиях низкого преваленса (как скрининговый), то оптимизации, наоборот, следует подвергать специфичность и прогностическую ценность отрицательного результата.

Разработано более 90 инструментов для оценки качества исследований диагностической точности [1]. Кокрановское сообщество в настоящее время рекомендует использование вопросника QUADAS, состоящего из 11 пунктов (<http://www.bristol.ac.uk/social-community-medicine/projects/quadas/>). Следует отметить, что валидизация данного вопросника не проводилась, однако он был разработан на основе формального консенсуса экспертов.

Мы выполнили перевод на русский язык (локализацию) и адаптацию вопросника QUADAS, а также предлагаем использовать шкалу на его основе, т. е. приписывать соответствующие баллы ответам по каждому вопросу с дальнейшим вычислением суммы баллов (табл. 1). Мы также предлагаем оценивать риск систематических ошибок как:

- низкий – при сумме баллов 8–11;
- средний – при сумме баллов 4–7;
- высокий – при сумме баллов 0–3.

Помимо включенных в вопросник 11 пунктов, при оценке отдельных исследований может возникнуть необходимость дополнительно оценить и другие аспекты исследования. Кокрановское сотрудничество приводит в [2] список таких возможных аспектов (табл. 2). При формулировании возникших дополнительных вопросов нужно следить за тем, чтобы ответ «да» на любой из их вопросов означал низкий риск систематической ошибки, а ответ «нет» – высокий.

Подчеркнем, что данный вопросник разработан для оценки одномоментных исследований. Так, в случае, если верификация теста проводится в отсроченный период (т. е. вместо использования референсного теста применяются результаты наблюдения), то исследование становится проспективным.

Зачастую выполнению оценки исследования диагностического теста препятствует недостаточно ясное описание выполненного исследования в публикациях [3, 4]. В связи с этим были разработаны рекомендации STARD (*stard-statement.org*) [5] по надлежащему представлению подобных исследований.

Общее суждение о методологическом качестве исследования диагностического теста рекомендуется выносить с учетом двух аспектов:

1. Риск систематических ошибок (может быть низким, средним, высоким).
2. Вероятность некорректности результатов статистического анализа (может быть низкой, средней, высокой), которая должна оцениваться экспертом с привлечением соответствующих методических материалов, например [6, 7].

Порядок вынесения общего суждения описан в нашей публикации [8].

ЛИТЕРАТУРА

1. Whiting P., Rutjes A. W., Dinnes J., Reitsma J. B., Bossuyt P. M., Kleijnen J. A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. *Journal of Clinical Epidemiology* 2005; 58: 1–12.

2. Reitsma J. B., Rutjes A. W. S., Whiting P., Vlassov V. V., Leeflang M. M. G., Deeks J. J. Chapter 9: Assessing methodological quality. In: Deeks J. J., Bossuyt P. M., Gatsonis C. (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. The Cochrane Collaboration, 2009. URL: <http://srdta.cochrane.org>.
3. Lumbreras-Lacarra B., Ramos-Rincón J. M., Hernández-Aguado I. Methodology in diagnostic laboratory test research in clinical chemistry and laboratory medicine. *Clinical Chemistry*. 2004; 50: 530–536.
4. Rutjes A. W., Reitsma J. B., Di Nisio M., Smidt N., van Rijn J. C., Bossuyt P. M. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; 174: 469–476.
5. Bossuyt P. M., Reitsma J. B., Bruns D. E., Gatsonis C. A., Glasziou P. P., Irwig L., Lijmer J. G., Moher D., Rennie D., de Vet H. C. W., Kressel H. Y., Rifai N., Golub R. M., Altman D. G., Hooft L., Korevaar D. A., Cohen J. F. For the STARD Group. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ*. 2015; 351: h5527. PMID: 26511519.
6. Ланг Т., Альтман Д. Основы описания статистического анализа в статьях, публикуемых в биомедицинских журналах. Руководство «Статистический анализ и методы в публикуемой литературе (САМПЛ)». Медицинские технологии. Оценка и выбор. 2014; 1(15), 11–16.
7. Реброва О. Ю. Описание статистического анализа данных в оригинальных статьях. Типичные ошибки. *Журнал неврологии и психиатрии им. С. С. Корсакова*. 2010; 110(11): 71–74.
8. Реброва О. Ю. О принципах экспертизы методологического качества оригинальных медицинских исследований. *Медицинские технологии. Оценка и выбор*. 2014; 4(18): 15–18.

REFERENCES

1. Whiting P., Rutjes A. W., Dinnes J., Reitsma J. B., Bossuyt P. M., Kleijnen J. A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. *Journal of Clinical Epidemiology* 2005; 58: 1–12.
2. Reitsma J. B., Rutjes A. W. S., Whiting P., Vlassov V. V., Leeflang M. M. G., Deeks J. J. Chapter 9: Assessing methodological quality. In: Deeks J. J., Bossuyt P. M., Gatsonis C. (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. The Cochrane Collaboration, 2009. URL: <http://srdta.cochrane.org>.
3. Lumbreras-Lacarra B., Ramos-Rincón J. M., Hernández-Aguado I. Methodology in diagnostic laboratory test research in clinical chemistry and laboratory medicine. *Clinical Chemistry*. 2004; 50: 530–536.
4. Rutjes A. W., Reitsma J. B., Di Nisio M., Smidt N., van Rijn J. C., Bossuyt P. M. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; 174: 469–476.
5. Bossuyt P. M., Reitsma J. B., Bruns D. E., Gatsonis C. A., Glasziou P. P., Irwig L., Lijmer J. G., Moher D., Rennie D., de Vet H. C. W., Kressel H. Y., Rifai N., Golub R. M., Altman D. G., Hooft L., Korevaar D. A., Cohen J. F. For the STARD Group. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ*. 2015; 351: h5527. PMID: 26511519.
6. Lang T., Altman D. Basic Statistical Reporting for Articles Published in Clinical Medical Journals: the SAMPL Guidelines. *Medical Technologies. Assessment and Choice*. 2014; 1(15): 11–16.

7. Rebrova O. Yu. Description of Statistical Analysis of Data in Original Articles. Typical Errors. *Zhurnal neurologii i psikiatrii imeni S. S. Korsakova*. 2010; 110(11): 71–74.
8. Rebrova O. Yu. On principles of assessment of medical research methodological quality. *Medical Technologies. Assessment and Choice*. 2014; 4(18): 15–18.

Сведения об авторах:

Реброва Ольга Юрьевна

профессор кафедры медицинской кибернетики и информатики РНИМУ им. Н. И. Пирогова, д-р мед. наук

Адрес для переписки:

117997, Москва, ул. Островитянова, д. 1

Тел.: +7 (495) 434-5478

E-mail: o.yu.rebrova@gmail.com

Федяева Влада Константиновна

главный специалист отдела методологического обеспечения проведения комплексной оценки технологий в здравоохранении ФГБУ ЦЭКМП Министерства здравоохранения РФ, старший лаборант Научно-исследовательского финансового института Минфина РФ

Адрес для переписки:

109028, Москва, Хохловский переулок, вл. 10, стр. 5.

Тел.: +7 (495) 783-1905

E-mail: vlada.fedyeva@gmail.com

Writing committee:

Rebrova Olga Yuryevna

Professor at the Department of Medical Cybernetics and Informatics of the N.I. Pirogov Russian National Research Medical University, PhD

Address for correspondence:

Ostrovityanova Str., 1, Moscow 117997, Russia

Tel.: +7 (495) 434-5478

E-mail: o.yu.rebrova@gmail.com

Fediaeva Vlada Konstantinovna

Chief specialist, Department of methodological support of comprehensive health technology assessment of the Center of Healthcare Quality Assessment and Control of Ministry of Health of the Russian Federation, senior assistant at the Scientific and Research Financial Institute, Ministry of Finances

Address for correspondence:

Khokhlovsky per., 10, bldg. 5, Moscow 109028

Tel.: +7 (495) 783-1905

E-mail: vlada.fedyeva@gmail.com