

Технологии консенсуса при экспертизе рекомендаций

И.Б. Заболотских¹, С.В. Григорьев¹, А.А. Белкин²,
Р.Е. Лакхин³

¹ ФГБОУ ВО «Кубанский государственный медицинский университет» Минздрава России, Краснодар, Россия

² ФГБОУ ВО «Уральский государственный медицинский университет» Минздрава России, Екатеринбург, Россия

³ ФГБВОУ ВО «Военно-медицинская академия им. С.М. Кирова» Министерства обороны России, Санкт-Петербург, Россия

Ключевые слова: экспертиза документов, AGREE II, метод Дельфи, метод номинальных групп, смешанный метод соответствия RAND, собрание по выработке консенсуса

✉ Для корреспонденции: Заболотских Игорь Борисович — д-р мед. наук, профессор, заведующий кафедрой анестезиологии, реаниматологии и трансфузиологии ФПК и ППС ФГБОУ ВО КубГМУ МЗ РФ, Краснодар, Россия; e-mail: pobeda_zib@mail.ru

✉ Для цитирования: И.Б. Заболотских, С.В. Григорьев, А.А. Белкин, Р.Е. Лакхин. Технологии консенсуса при экспертизе рекомендаций. Вестник интенсивной терапии им. А.И. Салтанова. 2020;4:8–11. DOI: 10.21320/1818-474X-2020-4-8-11

✉ Поступила: 08.11.2020

✉ Принята к печати: 16.11.2020

DOI: 10.21320/1818-474X-2020-4-8-11

В последние годы наблюдается нарастающее число экспертиз документов различного формата, в том числе и рекомендаций. Для этого используют ряд оценочных инструментов — как документа в целом, так и его отдельных положений, — в основе которых лежит достижение консенсуса (согласия) между экспертами.

Для экспертизы гайдлайна как целостного документа обычно используют модифицированный метод AGREE II (Appraisal of Guidelines for REsearch & Evaluation), позволяющий оценить соответствие конкретного рассматриваемого документа утвержденным формальным критериям и принципам экспертизы [1–3], которая включает 23 пункта, расположенных в 6 доменах [3]. В доменах оценивается актуальность (пункты 1–3), участие всех необходимых специалистов (пункты 4–6), соответствие формату и структуре гайд-

Consensus technologies in the analysis of guidelines

I.B. Zabolotskikh¹, S.V. Grigoryev¹, A.A. Belkin²,
R.E. Lakhin³

¹ Kuban State Medical University, Krasnodar, Russia

² Ural State Medical University, Yekaterinburg, Russia

³ S.M. Kirov Military Medical Academy

Keywords: document expertise, AGREE II, Delphi method, nominal group technique, RAND appropriateness method, NIH consensus conference

✉ For correspondence: Igor B. Zabolotskikh — Dr. Med. Sci., professor, head of Department of Anesthesiology, Intensive Care and Transfusiology, Kuban State Medical University; Krasnodar; Russia; e-mail: pobeda_zib@mail.ru

✉ For citation: I.B. Zabolotskikh, S.V. Grigoryev, A.A. Belkin, R.E. Lakhin. Consensus technologies in the analysis of guidelines. Annals of Critical Care. 2020;4:8–11. DOI: 10.21320/1818-474X-2020-4-8-11

✉ Received: 08.11.2020

✉ Accepted: 16.11.2020

лайна, наличие ключевых рекомендаций (пункты 15–17), возможность внедрения по выделенным критериям качества (пункты 18–21), конфликта интересов разработчиков (пункты 22–23). Третий домен является самым важным и трудоемким, так как включает проверку полноты поиска доказательных данных, уровней убедительности рекомендаций и достоверности доказательств, качества рецензирования и механизма просмотра гайдлайна (пункты 7–14).

После оценки каждого домена по 7-балльной шкале подводится общий рейтинг руководства в виде рекомендации о возможности или невозможности его применения в практической медицине. Оценка ближе к 1 баллу отражает степень несогласия, а ближе к 7 баллам — согласия. Преодоление порога в 5 баллов особенно важно для третьего домена.

Количество экспертиз в области анестезиологии и интенсивной терапии в базе PubMed, проведенных с помощью AGREE II, всего 12, а в общемедицинской практике — 723. Проанализируем итог экспертизы некоторых из них.

Оценка рекомендаций по аналгоседации у критических пациентов в ОРИТ [4] подразумевала анализ 5 гайдлайнов, и 4 эксперта оценили их качество как очень высокое — более 80 %. Экспертиза руководств по лечению хронической боли [5] включала 4 руководства, из них только у двух консенсусный порог в 60 % был преодолен, а другие два оказались низкого качества. 16 гайдлайнов были сгруппированы и проанализированы под общей тематикой оценки руководств по патологии минерального обмена и костей при ХБП [6], при этом 4 гайдлайна были согласованы, но на уровне минимально допустимого порога (более 60 %), остальные 12 — не преодолели эту планку. Рекомендации по применению игл для люмбальной пункции [9] получили наивысшую экспертную оценку — 93 %.

Важный и широкий пласт вопросов поднимается при экспертизе отдельных положений (или тезис-рекомендаций) руководств. Существует четыре наиболее распространенных метода формализованной оценки положений [10–15]:

- метод Дельфи (Delphi method);
- метод номинальных групп (Nominal Group Technique — NGT);
- смешанный метод соответствия RAND (RAND appropriateness method — RAM);
- собрание по выработке консенсуса (Consensus Conference NIH).

Суть *метода Дельфи* заключается в том, что координатор изначально формулирует обсуждаемую проблему либо для обсуждения берется готовый документ — например, гайдлайн или нормативно-правовой документ. В результате определения круга рассматриваемых проблем формулируют или выделяют положения, которые будут обсуждать эксперты. После проведения подготовительного этапа делается анонс экспертизы и приглашаются эксперты, которые знакомятся с анонсом и выражают согласие на участие. В случае согласия начинается работа с анкетами, то есть сама экспертиза. Дельфийский анализ подразумевает анонимное участие экспертов — эксперты не знают друг друга. Во время первого раунда каждый эксперт оценивает предложенные положения, при этом не только указывает степень своего согласия или несогласия с положением по балльной шкале, но и оставляет комментарий, почему он не согласен. Также эксперт может сформулировать свой, уточненный вариант положения.

После окончания первого раунда координатор суммирует сведения из анкет. Как правило, часть положений достигает порогового значения согласованности и считается принятой. Но ряд положений обычно нуждается в дальнейшей экспертизе, изменении

формулировок, их переоценке, в том числе с учетом комментариев экспертов — это может потребовать проведения второго, третьего и даже большего количества раундов. При этом все больше положений, как видоизмененных, так и вновь сформулированных в соответствии с замечаниями экспертов, достигает консенсуса. В итоге дельфийской экспертизы часть положений принимается в исходном виде, часть — в модифицированном, а ряд положений может быть отвергнут как не достигший консенсуса. Количество итоговых положений может соответствовать исходному, но чаще бывает больше или меньше исходного количества. Таким образом, дельфийский анализ демонстрирует активное аргументированное влияние экспертов на принятие итогового консенсусного соглашения.

С помощью дельфийской экспертизы достигнут консенсус по выбору дефиниций и целевых точек для периоперационной медицины [16]. В итоге согласованы 167 клинических индикаторов для прогнозирования клинических исходов.

Метод номинальных групп (Nominal Group Technique — NGT), в отличие от метода Дельфи, это очная экспертиза. Она включает 4 этапа. Количество экспертов небольшое — обычно рекомендуется до 7 человек, в отличие от дельфийского анализа, где в среднем число экспертов составляет около 15. Эксперты обсуждают положения, слышат друг друга, меняют свою точку зрения — в итоге вырабатывается единое согласованное мнение. В ходе 4 этапов исследования эксперты сначала знакомятся с проблемой за короткий промежуток времени (около 20 мин; так называемая «тихая генерация»). Как правило, обсуждаются не гайдлайны, а более узкие и конкретные проблемы. После этого происходит обсуждение вопросов — раунд «круглого стола». Эксперты могут уточнить свою точку зрения, сформулировать итоговые положения. После этого происходит этап голосования, с ранжированием мнений экспертов по каждому положению.

С помощью метода *NGT* проведена оценка подходов к поддержанию верхних дыхательных путей на догоспитальном этапе. Были согласованы оценка факторов риска для догоспитальной интубации, чек-лист и стандартные операционные процедуры, применение преоксигенации и препаратов для облегчения интубации [17].

Интересен «смешанный» *метод соответствия RAND* (RAND appropriateness method — RAM), разработанный корпорацией RAND (Research AND Development). Первой важной особенностью является то, что концепция «соответствия» подразумевает определение отношения «польза/риск» медицинских и хирургических процедур. Вторая особенность заключается в сочетании двух подходов: доказательного (на основе рандомизированных многоцентровых исследований и метаанализов) и коллективного мнения экспертов о применимости и эффективности процедур у боль-

шинства пациентов. Отсюда и название — смешанный метод соответствия.

Он включает в себя элементы двух ранее упомянутых методов — метода Дельфи и метода номинальных групп. От первого взят стартовый анонимный раунд, когда эксперты анонимно заполняют анкеты, в которых указаны рассматриваемые положения, и оценивают каждое положение в баллах, а второй раунд проходит очно, с обсуждением итогов первого раунда и составленных мнений по рассматриваемым проблемам. При необходимости достижения консенсуса по большому количеству положений можно провести дополнительный анонимный и очный раунд.

Этот метод был успешно применен при анализе рекомендаций по использованию ультразвука как у взрослых, так и у детей [16, 19]. Для взрослых из 24 рекомендаций были согласованы 12, касающихся применения ультразвука при плевральном выпоте и дренировании плевральной полости, пневмотораксе, катетеризации центральных вен и диагностике тромбоза глубоких вен. Для детей экспертиза согласовала 39 из 41 положения, касающегося применения ультразвука для оценки сердечно-сосудистой системы, легких, головного мозга, органов брюшной полости.

Собрание по выработке консенсуса Национального института здоровья США (National Institutes of Health Consensus Conference — NIH CC) подразумевает выбор для обсуждения потенциально решаемой проблемы. Вначале так называемые эксперты ядра, обычно от 10–12 человек, делают обзор литературы, формулируют положения и дают рекомендации по формированию списка «внешних» экспертов. На очной конференции собираются внешние эксперты, половина из которых обычно являются неспециалистами в рассматриваемой проблеме. Количество участников конференции, как правило, от нескольких сотен до тысячи человек, из которых каждый может высказаться по рассматриваемой проблеме. После проведения конференции эксперты ядра суммируют полученные результаты и формируют итоговый список положений.

С помощью метода *NIH CC* проведена оценка рекомендаций по трансфузии эритроцитов у детей в критическом состоянии [20]: было согласовано 102 положения, степень согласованности у всех была выше 80 %.

Порог консенсуса всегда устанавливается координатором до начала экспертной оценки. В основе определения порога консенсуса (обычно не менее 70 %), особенно применительно к методу Дельфи, смешанному методу соответствия RAND и технике AGREE II, лежит балльная оценка по одному из вариантов шкалы Лайкерта — от 3 до 10 баллов.

При поиске в базе данных PubMed по ключевым словам «anesthesia intensive care guidelines» обнаружено 2792 гайдлайна, из них только 88 подверглись экспертной оценке, что составляет 3,15 % от общего числа экспертиз в клинической медицине. Среди всех методов экспертиз, касающихся конкретных положений гайдлайнов, обращает на себя внимание наибольшая востребованность метода Дельфи, как максимально адекватного и ясного для объективной оценки любого формата документов. Смешанный метод соответствия RAND выгодно отличается разумным и иногда необходимым «очным этапом» взаимодействия экспертов. Предпочтительно сочетать один из этих методов с экспертизой документа как целого с помощью оценочного инструмента AGREE II (табл. 1).

Таким образом, в настоящее время имеется ряд специализированных инструментов, позволяющих оценить адекватность и качество как рекомендаций в целом, так и их отдельных положений. Участники любой экспертизы должны понимать, что их консенсусное мнение может влиять как на восприятие ранее опубликованных документов, особенно в части положений, не достигших порога консенсуса (ретроспективный аспект экспертизы), так и на разработку, изменение и утверждение новых технологий (проспективный аспект).

Настоящая редакторская заметка является анонсом подробного обзора по методу Дельфи, который будет опубликован в журнале «Вестник интенсивной терапии имени А.И. Салтанова», в первом номере за 2021 г., и последующих публикаций итогов проведенных экспертиз российских клинических и методических рекомендаций в области анестезиологии-реаниматологии.

ORCID авторов

Заболотских И.Б. — 0000-0002-3623-2546

Григорьев С.В. — 0000-0002-9753-7351

Белкин А.А. — 0000-0002-0544-1492

Лахин Р.Е. — 0000-0001-6819-9691

Таблица 1. Количество гайдлайнов, подвергнутых экспертной оценке разными методами

Table 1. Number of guidelines peer reviewed using different methods

Ключевые слова	Метод экспертной оценки				
	Delphi	NGT	RAND/RAM	NIH CC	AGREE II
«clinical practice guidelines»	1159	228	441/106	486	723
«anesthesia intensive care guidelines»	47	4	9/7	9	12

Литература/References

- [1] *Brouwers M.C., Kho M.E., Browman G.P., et al.* AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010; 182(18): E839–42.
- [2] *Brouwers M.C., Kho M.E., Browman G.P., et al.* Development of the AGREE II, part 1: performance, usefulness and areas for improvement. *CMAJ*. 2010; 182(10): 1045–1052.
- [3] *Brouwers M.C., Kho M.E., Browman G.P., et al.* Development of the AGREE II, part 2: assessment of validity of items and tools to support application. *CMAJ*. 2010; 182(10): E472–8.
- [4] *Girardis M., Cantaroni C., Savoia G., et al.* A critical appraisal of the quality of analgosedation guidelines in critically ill patients. *Minerva Anestesiologica*. 2016; 82(2): 230–235.
- [5] *Hoydonckx Y., Kumar P., Flamer D., et al.* Quality of chronic pain interventional treatment guidelines from pain societies: Assessment with the AGREE II instrument. *Eur J Pain*. 2020; 24(4): 704–721. DOI: 10.1002/ejp.1524
- [6] *Sekercioglu N., Al-Khalifah R., Ewusie J.E., et al.* A critical appraisal of chronic kidney disease mineral and bone disorders clinical practice guidelines using the AGREE II instrument. *Int Urol Nephrol*. 2017; 49(2): 273–284. DOI: 10.1007/s11255-016-1436-3
- [7] *Akyol M.U., Alden T.D., Amartino H., et al.* Recommendations for the management of MPS IVA: systematic evidence- and consensus-based guidance. *Orphanet J Rare Dis*. 2019; 14(1): 137. DOI: 10.1186/s13023-019-1074-9
- [8] *Akyol M.U., Alden T.D., Amartino H., et al.* Recommendations for the management of MPS VI: systematic evidence- and consensus-based guidance. *Orphanet J Rare Dis*. 2019; 14(1): 118. DOI: 10.1186/s13023-019-1080-y
- [9] *Rehn M., Chew M.S., Olkkola K.T., et al.* Clinical practice guideline on atraumatic (pencil-point) vs conventional needles for lumbar puncture: Endorsement by the Scandinavian Society of Anaesthesiology and Intensive Care Medicine. *Acta Anaesthesiologica Scandinavica*. 2019; 63(4): 438–439. DOI: 10.1111/aas.13312
- [10] *Dalkey N.C., Helmer O.* An experimental application of the Delphi method to the use of experts. *Management Science*. 1963; 9: 458–467.
- [11] *Delbecq A., Van de Ven A.* A group process model for problem identification and program planning. *Journal of Applied Behavioral Science*. 1971; 7: 467–492.
- [12] *Fink A., Kosecoff J., Chassin M., Brook R.H.* Consensus Methods: Characteristics and Guidelines for Use. *AJPH*. 1984; 74(9).
- [13] *McMillan S.S., King M., Tully M.P.* How to use the nominal group and Delphi techniques. *Int J Clin Pharm*. 2016; 38(3): 655–662. DOI: 10.1007/s11096-016-0257-x
- [14] *Diamond I.R. et al.* Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*. 2014; 67: 401e409.
- [15] *Linstone H.A., Turoff M.* The Delphi Method: Techniques and Applications. Addison-Wesley, 2002.
- [16] *Haller G., Bampoe S., Cook T., et al.* Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine initiative: clinical indicators // *Br J Anaesth*. 2019; 123(2): 228–237. DOI: 10.1016/j.bja.2019.04.041
- [17] *Sunde G.A., Kottmann A., Heltné J.K., et al.* Standardised data reporting from pre-hospital advanced airway management — a nominal group technique update of the Utstein-style airway template. *Scand J Trauma Resusc Emerg Med*. 2018; 26(1): 46. DOI: 10.1186/s13049-018-0509-y
- [18] *Frankel H.L., Kirkpatrick A.W., Elbarbary M., et al.* Guidelines for the Appropriate Use of Bedside General and Cardiac Ultrasonography in the Evaluation of Critically Ill Patients-Part I: General Ultrasonography. *Crit Care Med*. 2015; 43(11): 2479–2502. DOI: 10.1097/CCM.0000000000001216
- [19] *Singh Y., Tissot C., Fraga M.V., et al.* International evidence-based guidelines on Point of Care Ultrasound (POCUS) for critically ill neonates and children issued by the POCUS Working Group of the European Society of Paediatric and Neonatal Intensive Care (ESPNIC). *Crit Care*. 2020; 24(1): 65. DOI: 10.1186/s13054-020-2787-9
- [20] *Valentine S.L., Bembea M.M., Muszynski J.A., et al.* Consensus Recommendations for RBC Transfusion Practice in Critically Ill Children From the Pediatric Critical Care Transfusion and Anemia Expertise Initiative. *Pediatr Crit Care Med*. 2018; 19(9): 884–898. DOI: 10.1097/PCC.0000000000001613