The burden of conflicting guidelines

Olivier Steichen

The burden of conflicting guidelines

Olivier Steichen[a,b]

See original paper on page 2059

Evidence-based medicine (EBM) asserts that clinical decisions should be consistent with the results of clinical research. When performed from scratch, the full process involves [1]:

1. translating the clinical question into a generic question that can be answered by clinical research;
2. finding the studies that address this generic question;
3. appraising their quality and relevance;
4. critically synthesizing their results to answer the generic question; and
5. applying the generic answer to the clinical question with regards to the individual circumstances of the patient.

The full EBM process is time-consuming and requires technical skills from clinical epidemiology, information retrieval, and critical appraisal. EBM, therefore, recognizes that clinician should generally use evidence-based syntheses to guide their decisions rather than rely on original studies [2]. Systematic reviews attempt to provide the evidence-based answer to a focused generic question. But taking care of a patient requires the answer to many questions about diagnosis, prognosis, treatment, and follow-up. Finding and reading a relevant systematic review for each of these questions would remain very time-consuming. Moreover, clinicians would still have to decide if the answers provided by systematic reviews apply to their clinical settings and each of their patients.

Evidence-based guidelines go two steps further than systematic reviews:

1. they address all questions relevant to a given health problem or one of its broad aspects (e.g. diagnosis and prognosis, or treatment and follow-up);
2. they translate the evidence-based answers into recommendations, which are actionable decision rules.

The best guidelines rely on systematic reviews to find, appraise, and synthesize the evidence relevant to their topic. A diverse panel of contributors then interprets and translates the evidence into recommendations and specifies to which patients they should or might be applied, and how. Beyond evidence, guideline recommendations are expected to be sensitive to available resources, societal values, and individual preferences.

Good guidelines are very useful for those who have to make clinical decisions, although they do not master the intricacies of available evidence on a clinical topic. This is the case for junior physicians, still discovering their medical domain, and for primary care physicians, who can not know everything about everything. Guidelines are also invaluable resources for students, whose future depends on how well they perform at their assessment and are eager to have indisputable answers to test questions. For these enthusiastic guideline users, discordant recommendations understandably induce a feeling of helplessness and distrust.

Hundreds of clinical trials have been performed in the field of hypertension and one would expect that such an evidence base would leave little room for discordances. Unfortunately, discordant and even conflicting guidelines are the rule rather than the exception. A recent overview suggests that eight widely used guidelines on hypertension management in the general population, all published in English after 2010 and publicly available, disagree on 40% of their recommendations [3]. Adding to the confusion, conflicting guidelines sometimes arise from the same country, as for the 2014 Eighth Joint National Committee (JNC8), the 2017 American College of Cardiology/American Heart Association (ACC/AHA), and the 2017 American College of Physicians/American Academy of Family Physicians (ACP/AAFP) hypertension guidelines in the United States of America.

These disagreements can have profound consequences. Applied to the 2011–2014 US population, it has been estimated that [4]: antihypertensive drugs are indicated for 36.2% of adults according to the 2017 ACC/AHA guidelines, compared with 31.1% for the JNC8 (absolute difference: 11.5 million people); and 53.4% of adults receiving antihypertensive drugs are above goal blood pressure (BP) and require treatment intensification according to the 2017 ACC/AHA guidelines, compared with 24.7% for the JNC8 (absolute difference: 15.7 million hypertensive patients).

Discordances can arise from different guideline committees taking different evidences into account to support their

Journal of Hypertension 2020, 38:1945–1947

[a]Department of internal medicine, Tenon hospital, Assistance Publique – Hôpitaux de Paris and [b]LIMICS, UMRS 1142, Inserm, Sorbonne Université, Université Paris, Paris, France

Correspondence to Olivier Steichen, Department of Internal Medicine, Hôpital Tenon, 4 rue de la Chine, 75020 Paris, France. Tel: +33 1 56 01 78 31; fax: +33 1 56 01 71 46; e-mail: olivier.steichen@aphp.fr

DOI:10.1097/HJH.0000000000002549

recommendations. This is especially true when there is no formal process to ensure that the recommendations are evidence-based. In these cases, one can fear that selected evidence is used to support the recommendations endorsed by the guideline committee, rather than all available evidence retrieved and critically appraised to derive the recommendations. A study examined four trusted hypertension guidelines issued in English between 2003 and 2006 [5]. The definition of hypertension and the general organization of the guidelines were highly similar, but there were major differences in the detailed content and recommendations. The evidence used to support each guideline was then analysed: 350–550 references were cited per guideline, totalling 1402 references overall. Only 17 (1.2%) references were common to all four guidelines and 30 (2.1%) common to three guidelines. The overwhelming majority of references (1941, 88%) were cited by a single guideline. These figures suggest that confirmatory evidence was selected by each committee to support consensus-based guidelines rather than the whole body of available evidence comprehensively reviewed to derive evidence-based guidelines.

Over the last 20 years, however, guideline development has moved toward a more rigorous process. The EBM movement has again been engaged in these advances. The GRADE framework is a systematic approach to produce and present evidence-based clinical practice guidelines [6]. Its main objective is to provide a transparent and structured assessment of the strength of each recommendation and the quality of supporting evidence. In parallel, the AGREE (Appraisal of Guidelines for Research and Evaluation) Enterprise provides tools to evaluate the development process and the reporting of practice guidelines. The AGREE II tool lists 23 items into six domains [7]. It has been thoroughly validated and extensively used for the comparative appraisal of guidelines, including in the field of hypertension [8–10].

The authors of a study published in this issue identified eight high-quality hypertension guidelines (AGREE II score at least 60% for the 'rigour of development' and 'editorial independence' domains), written in English and published from 2016 to 2019 [11]. Then, they looked for conflicting treatment recommendations and discussed the implications for local adoption or adaptation. The topics covered by the guidelines were similar: lifestyle changes; indications of pharmacotherapy; choice of pharmacotherapy; therapeutic goals; adherence strategies; consideration of special patient groups. Recommendations were much more concordant between these rigorously developed guidelines than between less selected hypertension guidelines [3]. Nonetheless, discordances remained about the BP threshold to initiate pharmacotherapy and the BP target to reach, especially in young patients (low short-term cardiovascular risk but long life expectancy) and elderly patients (high short-term cardiovascular risk but high risk of side effects and shorter life expectancy).

The conflicting recommendations were often rated as strong and evidence-based by their respective writing committee. How can recommendations conflict if they are based on a thorough review of available evidence? As Mancia and Zanchetti [12] pointed out 20 years ago, evidence-based practice is not the direct application of objective evidence but results from the interpretation and extrapolation of this evidence. The importance of interpretation and extrapolation has been increasingly recognized by EBM. Guideline committees are expected to engage in a deliberative process to weight, which research is the most valid, relevant, and applicable, and the relative importance of anticipated benefits, risks, and costs [13]. Different committees may have valid reasons to disagree on these matters [14]. They can also have less valid reasons to disagree, such as conflicts of interest, disregard of relevant evidence, inappropriate weight given to relevant evidence, disregard of relevant outcomes, inappropriate weight given to relevant outcomes. The authors of the present study convincingly show that conflicting recommendations from high-quality hypertension guidelines indeed result from diverse interpretations, extrapolations and prioritization of a few key studies. A typical example is the SPRINT trial [15], which implications for the BP target to reach with antihypertensive treatment are still debated.

Another point raised by the authors is that all these high-quality guidelines scored low in the applicability domain of the AGREE II instrument. This domain covers four points [7]:

1. Does the guideline discuss facilitators and barriers to the recommendations? For example, do the potential users have the knowledge and skills to deliver the recommended care?
2. Does the guideline provide advice or tools on how the recommendations can be applied in practice? For example, are summary documents, checklists, or algorithms provided?
3. Does the guideline discuss the resource needed to apply the recommendations? For example, are recommended drugs and investigations affordable?
4. Does the guideline provide criteria to monitor the application of recommendations? For example, are criteria proposed to measure the adherence to recommendations or their impact on the results of care?

Improving the applicability of high-quality guidelines is naturally important to facilitate local adoption and adaptation. But identifying conflicts between recommendations and understanding their cause is also important to this end. It should be clear when the body of available evidence allows diverging interpretations and extrapolations. In these cases, guideline committees should avoid strong recommendations and leave room for flexibility and clinical autonomy [16,17]. This freedom must be used to individualize clinical decisions, when deciding whether and how to apply recommendations to the unique circumstances and preferences of a given patient. And it should also be used to adapt the guidelines to specific settings (healthcare facility, administrative region, country), when deciding if the recommendations fit the community values and behaviours, and which resources are available for their implementation.

The applicability domain of the AGREE II tool only scratches the surface of issues related to the application of recommendations to diverse patients [18,19] and diverse clinical settings [20,21]. As EBM acknowledges it, clinical

research evidence is only a (small) part of reasonable medical decision-making [22,23]. Recommendations should not be viewed as commandments and guidelines should help to prioritize them according to clinical circumstances, values and preferences, and available resources [21,24].

ACKNOWLEDGEMENTS

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *BMJ* 1995; 310:1122–1126.
2. Guyatt GH, Meade MO, Jaeschke RZ, Cook DJ, Haynes RB. Practitioners of evidence based care. Not all clinicians need to appraise evidence from scratch but all need some skills. *BMJ* 2000; 320:954–955.
3. Alper BS, Price A, van Zuuren EJ, Fedorowicz Z, Shaughnessy AF, Oettgen P, *et al.* Consistency of recommendations for evaluation and management of hypertension. *JAMA Netw Open* 2019; 2:e1915975.
4. Muntner P, Carey RM, Gidding S, Jones DW, Taler SJ, Wright JT, *et al.* Potential U. S. population impact of the 2017 ACC/AHA high blood pressure guideline. *J Am Coll Cardiol* 2018; 71:109–118.
5. Georg G, Colombet I, Durieux P, Ménard J, Meneton P. A comparative analysis of four clinical guidelines for hypertension management. *J Hum Hypertens* 2008; 22:829–837.
6. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, *et al.*, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336:924–926.
7. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, *et al.* AGREE II: advancing guideline development, reporting and evaluation in healthcare. *J Clin Epidemiol* 2010; 63:1308–1311.
8. Al-Ansary LA, Tricco AC, Adi Y, Bawazeer G, Perrier L, Al-Ghonaim M, *et al.* A systematic review of recent clinical practice guidelines on the diagnosis, assessment and management of hypertension. *PloS One* 2013; 8:e53744.
9. Zhang X, Zhao K, Bai Z, Yu J, Bai F. Clinical practice guidelines for hypertension: evaluation of quality using the AGREE II Instrument. *Am J Cardiovasc Drugs* 2016; 16:439–451.
10. Okwen PM, Maweu I, Grimmer K, Margarita Dizon J. Evaluation of all African clinical practice guidelines for hypertension: quality and opportunities for improvement. *J Eval Clin Pract* 2019; 25:565–574.
11. Leite-Santos NC, de Melo DO, Mantovani-Silva RA, Gabriel FC, Fornasari GS, Dórea EL, *et al.* Guidelines for hypertension management in primary care: is local adaptation possible? *J Hypertens* 2020; 38:2059–2073.
12. Mancia G, Zanchetti A. Evidence-based medicine: an educational instrument or a standard for implementation? *J Hypertens* 1999; 17:1509–1510.
13. Djulbegovic B, Guyatt G. Evidence vs consensus in clinical practice guidelines. *JAMA* 2019; doi:10.1001/jama.2019.9751 [Epub ahead of print].
14. Oxman AD, Glasziou P, Williams JW Jr. What should clinicians do when faced with conflicting recommendations? *BMJ* 2008; 337:a2530.
15. SPRINT Research Group: Wright JT, Williamson JD, Whelton PK, Snyder JK, Sink KM, *et al.* A randomized trial of intensive versus standard blood-pressure control. *N Engl J Med* 2015; 373:2103–2116.
16. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, *et al.*, GRADE Working Group. Going from evidence to recommendations. *BMJ* 2008; 336:1049–1051.
17. Kastner M, Bhattacharyya O, Hayden L, Makarski J, Estey E, Durocher L, *et al.* Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review. *J Clin Epidemiol* 2015; 68:498–509.
18. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005; 294:716–724.
19. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Res Policy Syst* 2006; 4:22.
20. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Res Policy Syst* 2006; 4:25.
21. Brouwers MC, Makarski J, Kastner M, Hayden L, Bhattacharyya O, GUIDE-M Research Team. The Guideline Implementability Decision Excellence Model (GUIDE-M): a mixed methods approach to create an international resource to advance the practice guideline field. *Implement Sci* 2015; 10:36.
22. Montori VM, Guyatt GH. Progress in evidence-based medicine. *JAMA* 2008; 300:1814–1816.
23. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet* 2017; 390:415–423.
24. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, *et al.*, el GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: clinical practice guidelines. *BMJ* 2016; 353:i2089.