

Guidelines: we'll always need them, we sometimes dislike them, and we have to make them better

R. Kahn

Received: 8 July 2010 / Accepted: 28 July 2010 / Published online: 25 August 2010
© Springer-Verlag 2010

Keywords Clinical guidelines · Diabetes guidelines · Guidelines

Not long ago clinical decisions were mostly based on what someone's colleagues were doing, or upon the personal experience of a well-thought-of clinician as conveyed during a lecture or in a hallway discussion. There was no written rationale, no review of published studies and no distillation of the evidence that systematically presented the advantages and disadvantages of a treatment. All this fostered enormous variations in practice, much inappropriate care, misuse, overuse and underuse of medical services, and considerable physician uncertainty about how best to treat disease.

The value of guidelines

Although all of the above problems are still worrisome, there is no question that the advent of clinical guidelines has helped enormously in reducing these and other problems [1–5]. For example, guidelines have become critically important tools that help simplify the complexity of medicine. They free clinicians from spending huge amounts of time sorting through and grasping the results (and details) of innumerable research publications on a given subject in order to arrive at a course of action. They provide reassurance that the 'right' course of action has been recommended by experts in the field, who themselves have distilled the relevant data, pooled their expertise and applied appropriate judgement to a medical condition.

The way guidelines have become the cornerstone of medicine is exemplified by the fact that they: (1) are essential teaching tools for medical students (e.g. textbooks); (2) appear in drug labelling (e.g. indications and contraindications); (3) are pre-requisite knowledge for medical licensure (e.g. the questions on board exams); (4) serve as basis for the evaluation and improvement of quality care in a population (e.g. performance measures); and (5) often function as the basis for reimbursement (e.g. Medicare, National Institute for Health and Clinical Excellence [NICE]). Guidelines are now so numerous and continually being written and updated by a wide range of medically related associations, medical specialty societies, managed care organisations and other groups, that they often trip over each other [6]. Although it has been difficult to draw an unequivocal straight line from a specific guideline to a change in the practice of medicine, it's hard to imagine, for example, that the way diabetes care is delivered, from screening and detection through to diagnosis and ongoing medical management, not to mention the progress made in past decades is not the result of diabetes guidelines published in the last 20 or 30 years.

Some examples

Two recently published guidelines exemplify both the value and, as we'll see, the shortcomings of the enterprise [7, 8]. The need for the first one [7] stems from the conflicting results of large, randomised control studies that addressed the effect of glycaemic control on incidence of CVD. After the Diabetes Control and Complications Trial (DCCT) and UK Prospective Diabetes Study (UKPDS) had shown the unequivocal benefit of glucose-lowering for preventing microvascular complications of diabetes, there remained great uncertainty whether intensive therapy lowered the risk of CVD, which is arguably the most impactful complication

R. Kahn (✉)
Department of Medicine, University of North Carolina,
Chapel Hill, NC 27599, USA
e-mail: rak6200@gmail.com

associated with diabetes. Subsequent long-term follow-up of participants in both studies indicated that patients undergoing intensive therapy experienced a significant reduction of CVD events [9, 10], suggesting that excellent glucose control could reduce the incidence of macrovascular complications. Just when it all seemed clear and straightforward, along came the results of three other trials specifically designed to address the benefits and risk of intensive therapy with regard to cardiovascular events [11–13]. Surprisingly, all three showed no beneficial effect of intensive glucose-lowering on a composite of CVD endpoints. Indeed, one (the Action to Control Cardiovascular Risk in Diabetes [ACCORD] trial [11]) was terminated early because of excess deaths in the intensively treated group. Moreover, all three studies reported a significantly increased frequency of serious hypoglycaemic events related to intensive therapy. Subset analyses in all three studies subsequently provided information that might be useful for interpreting the main results and also shed light on which, if any, patients might benefit from intensive control. Of note, in all five studies, there were many significant differences in the patient population recruited, the therapy administered, HbA_{1c} levels achieved and in the outcomes tracked.

How should a clinician interpret these studies? Should intensive therapy be offered to no one, everyone or only some subset(s) of the diabetic population? If there ever was a purpose and value of a guideline, this subject would seem to beg for one that could bring order, direction and rationality to an important topic.

Recognising the confusing nature of these studies, three prominent organisations appointed a group of experts in diabetes and CVD to review the data and provide treatment recommendations. The resulting guideline [7] reviews the benefits and risks of intensive therapy, discusses the nuances of the studies and summarises what it all means and the implications for clinical care. Precisely what a guideline is suppose to do.

The second example [8] relates to the complicated subject of the pharmacological treatment of hyperglycaemia in diabetes. With nine classes of glucose-lowering agents, at least that many pharmaceutical companies each spending many millions of dollars advertising the virtues of their ‘unique’ product and a literature base notable for its lack of many well controlled trials directly comparing different agents, it’s easy to appreciate the confusion of many clinicians when it comes to deciding what to use when starting therapy, what to use when the initial therapy fails, and how and when to use combination therapy. Should generalists wade through the umpteen papers published in more than two dozen journals, each reporting the benefit of one or two of the drugs, and then attempt themselves to devise a treatment algorithm? That’s unlikely to happen.

To provide some guidance in this area, a writing group was given a ream of publications related to the pharmacological treatment of diabetes, along with the specific publications suggested by each manufacturer of an available drug. Altogether there were dozens of articles to discuss and debate. In the end, the group developed a simple algorithm that highlighted the use of some agents and downplayed the use of others. One agent was not recommended. Although, undoubtedly, the guideline could have been much more complex, detailed or nuanced, the writing group felt their recommendations provided sufficient guidance to practitioners who cannot devote resources or time to remembering all the benefits, drawbacks and special characteristics of a wide variety of agents, each with virtually the same indication. Here, once again, is a guideline intended to reduce the complexity of medical decisions and free clinicians from estimating and weighing the pros and cons of a panoply of drugs.

There is more to a guideline than the evidence

In the case of the two guidelines cited above, other authors arrived at different conclusions in their review of virtually the same evidence base [14–16]. Although some criticism of guidelines is rightfully directed toward conflicts of interest in members of the writing panels or the failure to cover some of the important elements of a guideline [17–23], that did not seem to be the predominant reason for the alternative views. Indeed, it’s certainly possible that, even when major shortcomings are revealed in a guideline, it is still possible to reach the very same conclusions as those of the ‘flawed’ guideline [24]. This is because critics of guidelines in general [20–22] or in specific cases [16] almost always overlook or ignore the fact that a major component of all guidelines is judgement.

In the first of a comprehensive, landmark series of papers on clinical and health policy decision making (all authored by D. M. Eddy and published in the *Journal of the American Medical Association* in 1990), it was pointed out that scientific evidence and judgement are the two key components [25] (Fig. 1) of a medical decision or policy. No doubt much criticism and debate around (a) guideline(s) results from the failure to appreciate that at every step of guideline development judgement is required. Judgement starts with the evaluation of the very clinical studies that form the basis of the guideline. For example, were the inclusion or exclusion rules too loose or rigid, was the drop-out rate too high, were the outcomes measured reasonable and fully documented, how well or properly was the trial executed, to what extent were the results clinically meaningful? Not surprisingly, therefore, the results of major, randomised control trials are routinely commented on in journal editorials, which point out strengths and weaknesses of the study in question,

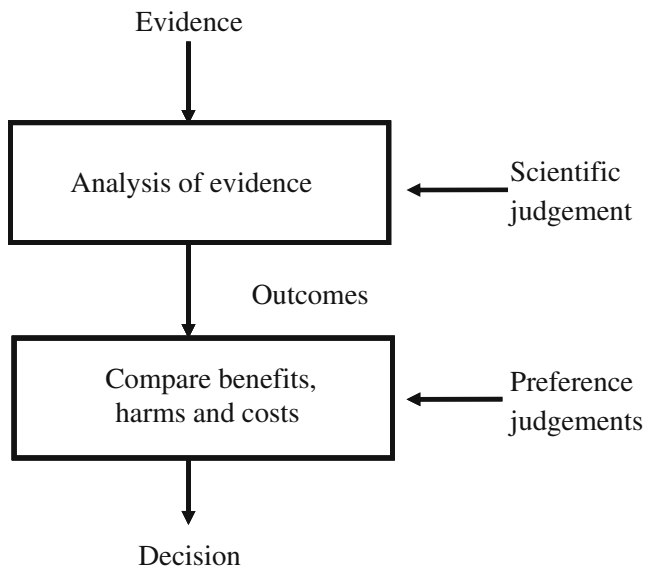


Fig. 1 Components of a medical decision. Adapted with permission from Eddy [25]

highlighting the fact that even the adequacy of the evidence is a matter of judgement or opinion, not an objective characteristic of the evidence itself [26]. Of course, the benefits, harms and cost of any treatment have trade-offs, the evaluation of which is especially subjective when interventions are compared and different experts bring their own opinions and biases to bear. Finally, judgement plays a key role in evaluating the weight of the evidence, especially when studies are combined [27]. Clearly, therefore, even the best of clinical trials or a summary of them do not necessarily provide clear-cut guidance. Evidence requires interpretation, which in turn requires judgement, which in turn requires the expression of opinion.

Thus, no guideline, however much it stems from randomised controlled trials, is purely or even largely ‘evidence-based’, if that term is meant to imply that the recommendations were not greatly influenced by opinion, values and personal preferences, i.e. by judgement. To be sure, the subjective elements are derived from a platform of science, but they themselves are not scientific judgements [26]. It follows that papers criticising a guideline or even the construction of guidelines should be very clear about whether the disagreement or problem is about science (e.g. does evidence exist) or judgement (e.g. what does it mean).

What does this imply?

Does the subjectivity of guidelines imply we should do away with them? Not at all. As pointed out above, guidelines serve a very useful purpose. When experts review the evidence and come to a consensus on what it means (all with great transparency), physicians are provided

with an informed opinion that obviates a lot of uncertainty, among other benefits. It would be nice if the consensus were forged by hundreds of opinions instead of those of a small writing group, but so far that seems difficult to implement and may be especially problematic if (as there always will be) some ‘experts’ refuse to compromise. Their reservations may of course be ‘right’, but whether practitioners would take the time and effort to study the various perspectives remains to be seen.

Other issues

Even under the best of circumstances, in the most obvious and agreed-upon guidelines, there will invariably be other problems, some inherent or fundamental, and these will always limit the utility of a guideline.

First, no set of trials or evidence can possibly address all the known caveats relevant to a clinical decision. In diabetes, patients can be at different stages of the disease process, with varying clinical manifestations of the diabetes syndrome. Outcomes in randomised trials represent averages derived from selected groups of patients, and the benefits or an intervention depend on many patient factors, most notably age, life expectancy and co-morbidities [28]. It is the complexity of these and other factors that by necessity makes the guidelines cited above [7, 8] so nuanced, because it is impossible to standardise therapy for the wide variety of patients seen in daily clinical practice. In the case of the pharmacotherapy guideline [8], critics might then say something like ‘so every drug is good for someone’, which in fact the guideline acknowledges, but the writing group still believed it important to define a preferred route of therapy. In their judgement of the scientific evidence along with other clinical judgements deemed relevant (e.g. cost), they arrived at what they thought to be the best established, most effective and most cost-effective strategy for achieving target glycaemic goals, albeit with an algorithm that still leaves much room for clinician discretion.

Second, virtually all guidelines do not formally address the cost-effectiveness of the intervention(s) recommended. That is not necessarily an omission by design; rather, there are usually few relevant studies focusing on this point. The impact of costs in medicine has only recently been an issue, and for the vast majority of interventions for which guidelines exist, cost is rarely acknowledged let alone discussed, because studies on this aspect simply don’t exist. However, omitting the cost factor makes it impossible to arrive at the real value of an intervention. Everything offering benefit or relative benefit is recommended, yet it’s seldom acknowledged that a given intervention known to improve health outcomes may in fact be very expensive and of low value. In a recent report examining the cost-

effectiveness of various approaches to preventing CVD, the cost-effectiveness of 11 commonly recommended/endorsed interventions differed by up to 100-fold [29]!

Finally, current guidelines do not always compare the full range of options available to treat a problem; yet if the full range is considered, this is likely to result in quite different benefits [29]. Also, when dealing with continuous variables (e.g. blood pressure, lipid and glucose levels), guidelines usually establish artificial cutpoints for when to begin treatment. For example, a number of interventions are available to reduce cardiovascular disease, yet no guideline weighs the comparative merits of blood pressure-, lipid- and glucose-lowering in light of a patient's family and medical history along with the pre-treatment laboratory values for each of those variables. Instead, again by necessity, a guideline will usually say 'treat X if it's above Y' and not deal with other important variables, or if it does, then only superficially. Although on occasion a guideline will provide one or two other trigger points depending on other factors, it cannot possibly cover the entire waterfront of circumstances known to be relevant. Alternatively, as demonstrated in the two examples above, the recommendations will have many caveats and will be proscriptive only to a modest degree. In short, although helpful, they are far from definitive.

Can it be done better?

This paper is certainly not the first to outline some of the shortcomings of guidelines. Moreover, as pointed out earlier, many authors have suggested important technical remedies like better grading systems and the elimination of obvious conflicts. Others have recommended more fundamental remedies [30], which while important and helpful, will not and cannot lead to a recommended course of action that considers the complete profile of the patient. Some have seemed to try, such as with the recommendations for management of ST-elevation myocardial infarction [31], but with over 400 recommendations and some 70 tables and figures, such documents are not very accessible to clinicians and may still be incomplete. Programming such documents into electronic medical records will not be the way to weigh and prioritise treatment based on all the options available and all the relevant patient characteristics. Others pin their hopes on critically needed research that will hopefully uncover genotypes or biomarkers enabling disease subtypes and the way they respond to therapy to be identified [32]. At present, however, there is no evidence that this approach will result in meaningful algorithms that can be used to guide therapy in more than a fraction of cases. More recently, Hayward and colleagues [33] developed an individualised approach to statin therapy, but their

tool for assessing risk and guiding treatment (the Framingham equation) is clinically simplistic and was not designed, nor has it been validated, to make specific treatment decisions. Their observation, however, that using even a crudely tailored treatment approach seems superior to a treat-to-target strategy indicates that individualised guidelines have enormous potential.

One approach of great promise is the use of computer-generated algorithms to produce patient-specific guidelines. Comprehensive, well-validated mathematical models of disease and healthcare delivery exist [34] and are already being used to predict the risk of disease [35] and the likelihood of benefit from one of many different treatments [36], depending on a person's demographic and medical profile. It would not seem an enormous leap for such software to produce individualised patient guidelines that take the next step, namely to weigh and prioritise a wide variety of treatments, although such an approach has yet to be demonstrated and validated.

In the meantime, we are stuck with what is necessary but inadequate: our current clinical guidelines. We need them, we know they pose problems, and we must find a way to replace them with something better.

Duality of interest The author declares that there is no duality of interest associated with this manuscript.

References

1. Eddy DM (1982) Clinical policies and the quality of clinical practice. *N Engl J Med* 307:343–347
2. Eddy DM (1990) Practice policies—what are they? *JAMA* 263:877–880
3. Field MJ, Lohr FN (eds) (1992) Guidelines for clinical practice. From development to use. Institute for Medicine committee on clinical practice guidelines. National Academy Press, Washington
4. Bero LA, Grilli R, Grimshaw JM et al (1998) Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. *BMJ* 317:465–468
5. Corrigan JM, Burstin H (2007) Measuring quality of performance: where it is headed and who is making the decisions? *J Fam Pract* 56:4A–7A
6. Kahn RA, Gale EM (2010) Gridlocked guidelines for diabetes. *Lancet* 375:2203–2204
7. American Diabetes Association, American College of Cardiology Foundation, American Heart Association (2009) Intensive glycemic control and the prevention of cardiovascular events: implications of the ACCORD, ADVANCE, and VA Diabetes Trials. *Diabetes Care* 32:187–192
8. American Diabetes Association, European Association for the Study of Diabetes (2009) Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy. *Diab Care* 32:193–203
9. Nathan DM, Cleary PA, Backlund JY et al (2005) Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med* 353:2643–2653

10. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA (2008) 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med* 359:1577–1589
11. Action to Control Cardiovascular Risk in Diabetes Study Group (2008) Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 358:2545–2559
12. Group AC, Patel A, MacMahon S et al (2008) Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med* 358:2560–2572
13. Duckworth W, Abraira C, Moritz T et al (2009) Glucose control and vascular complications in veterans with type 2 diabetes. *N Engl J Med* 360:129–139
14. Lehman R, Krumholz HM (2009) Tight control of blood glucose in longstanding type 2 diabetes. *BMJ* 338:901–902
15. Havas S (2009) The ACCORD trial and control of blood glucose level in type 2 diabetes mellitus. *Arch Intern Med* 169:150–154
16. Schernthaner G, Barnett AH, Betteridge DJ et al (2010) Is the ADA/EASD algorithm for the management of type 2 diabetes (January 2009) based on evidence or opinion? A critical analysis. *Diabetologia* 53:1258–1269
17. Woolf SH (1993) Practice guidelines: a new reality in medicine. III. Impact on patient care. *Arch Intern Med* 153:2646–2655
18. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J (1999) Are guidelines following guidelines? The methodological quality of clinical practice guidelines published in the peer-reviewed medical literature. *JAMA* 281:1900–1905
19. Choudhry NK, Stelfox HT, Detsky AS (2002) Relationship between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 287:612–617
20. Detsky AS (2006) Sources of bias for authors of clinical practice guidelines. *CMAJ* 175:1033–1035
21. McAlister FA, van Diepen S, Padwal RS et al (2007) How evidence based are the recommendations in evidence-based guidelines. *PLoS Med* 4:1325–1332
22. Sniderman AD, Furberg CD (2009) Why guideline-making needs reform. *JAMA* 301:429–431
23. Shiffman RN, Shekelle P, Overhage M et al (2003) Standardized reporting of clinical practice guidelines: a proposal from the conference on guideline standardization. *Ann Intern Med* 139:493–498
24. Qasseem A, Vijan S, Snow V (2007) Glycemic control and type 2 diabetes mellitus: the optimal hemoglobin A1c targets. A guidance statement from the American College of Physicians. *Ann Intern Med* 147:417–422
25. Eddy DM (1990) Anatomy of a decision. *JAMA* 263:441–443
26. Tunis SR (2007) Reflections on science, judgement, and value on evidence-based decision making: a conversation with David Eddy. *Health Aff* 26:w500–w515
27. Shrier I, Boivin J-F, Platt RW et al (2008) The interpretation of systematic reviews with meta-analyses: an objective or subjective process? *BMC Med Inform Dec Making* 8:19
28. Boyd CM, Darer J, Boult C (2005) Clinical practice guidelines and the quality of care for older patients with multiple comorbid diseases. *JAMA* 294:716–724
29. Kahn R, Robertson RM, Smith R, Eddy D (2008) The impact of prevention on reducing the burden of cardiovascular disease. *Circulation* 118:576–585
30. O'Connor PJ (2005) Adding value to evidence-based clinical guidelines. *JAMA* 294:741–743
31. Antman EM, Anbe DT, Armstrong PW et al (2004) ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 110:e82–e292
32. Smith RJ, Nathan DM, Arslanian SA et al (2010) Individualizing therapies in type 2 diabetes mellitus based on patient characteristics; what we know and what we need to know. *J Clin Endocrinol Metab* 95:1566–1574
33. Hayward RA, Krumholz HM, Zulman DM, Timble JW, Vijan S (2010) Optimizing statin treatment for primary prevention of coronary artery disease. *Ann Intern Med* 152:69–77
34. Eddy DM, Schlessinger L (2003) Archimedes: a trial validated model of diabetes. *Diab Care* 26:3093–3101
35. Stern M, Williams K, Eddy DM, Kahn R (2008) Validation of prediction of diabetes by Archimedes and comparison with other predicting models. *Diab Care* 31:1670–1671
36. American Diabetes Association. Diabetes PHD. Available from www.diabetes.org/living-with-diabetes/complications/diabetes-phd/. Accessed 17 June 2010