Progress in evidence-based medicine: a quarter century on

Benjamin Djulbegovic, Gordon H Guyatt

In response to limitations in the understanding and use of published evidence, evidence-based medicine (EBM) began as a movement in the early 1990s. EBM's initial focus was on educating clinicians in the understanding and use of published literature to optimise clinical care, including the science of systematic reviews. EBM progressed to recognise limitations of evidence alone, and has increasingly stressed the need to combine critical appraisal of the evidence with patient's values and preferences through shared decision making. In another progress, EBM incorporated and further developed the science of producing trustworthy clinical practice guidelines pioneered by investigators in the 1980s. EBM's enduring contributions to clinical medicine include placing the practice of medicine on a solid scientific basis, the development of more sophisticated hierarchies of evidence, the recognition of the crucial role of patient values and preferences in clinical decision making, and the development of the methodology for generating trustworthy recommendations.

Published **Online**
February 16, 2017
http://dx.doi.org/10.1016/S0140-6736(16)31592-6

University of South Florida Program for Comparative Effectiveness Research, and Division of Evidence Based Medicine, Department of Internal Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, USA (Prof B Djulbegovic MD); **H Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA** (Prof B Djulbegovic); **Tampa General Hospital, Tampa, FL, USA** (Prof B Djulbegovic); **and Department of Clinical Epidemiology and Biostatistics, and Department of Medicine, McMaster University, Hamilton, ON, Canada** (Prof G H Guyatt MD)

Correspondence to:
Prof Benjamin Djulbegovic, USF Health, Program for Comparative Effectiveness Research, University of South Florida, Tampa, FL 33612, USA
bdjulbeg@health.usf.edu

Historical origins of evidence-based medicine

Since the time of Hippocrates, medicine has struggled to balance the uncontrolled experience of healers with observations obtained by rigorous investigation of claims regarding the effects of health interventions. During the past 300 years, demands that the practice of medicine be founded on scientifically trustworthy empirical evidence have become increasingly vocal.

Pioneers, including Rudolph Virchow, Claude Bernard, and Louis Pasteur, championed science in medicine in Europe, and the Flexner report in the early 20th century cemented scientific inquiry as a bedrock of American medicine. Although one can identify attempts to obtain accurate observational data in the work of Pierre-Charles-Alexandre Louis and John Snow in the mid-19th century, and the use of clinical trials in James Lind's famous study of scurvy in the British navy,[1] the focus of most of these innovators was on physiological and basic research as a foundation for clinical practice, rather than the empirical assessment of diagnostic testing, prognosis, and therapeutic effect.

Indeed, it was not until 1962, with the passage of the US Food and Drug Administration Kefauver-Harris Act in the USA, that rigorous empirical testing of clinical trials in human beings was legally required to establish claims regarding drug efficacy;[2] other countries followed soon thereafter. Although these regulatory developments established the necessity for well done clinical trials demonstrating the efficacy and safety of new pharmaceutical innovations, unsystematic, uncontrolled clinical experience and physiological reasoning maintained their dominance as drivers of clinical practice.

In the 1970s and 1980s, David Sackett, David Eddy, and Archie Cochrane (among others) highlighted the need for strengthening the empirical practice of medicine and proposed initial evidentiary rules for guiding clinical decisions.[3–7] In 1991, one of us (GHG) introduced the term evidence-based medicine (EBM),[8] with a focus on educating front-line clinicians in assessing the credibility of research evidence,

Key messages

- EBM started as a movement in the early 1990s to evaluate and in turn acquire a better empirical basis for the practice of medicine
- EBM originally focused on critical appraisal, development of systematic reviews, and clinical practice guidelines
- These three domains coalesced in the mid-2000s and characterise the practice of EBM today
- EBM has become essential for the training of young clinicians by stressing critical thinking and the importance of statistical reasoning and continuous evaluation of medical practice
- EBM has contributed substantially to improvement of the quality of research by transparently documenting the problems with existing research and subsequently developing better research standards
- EBM has also improved the practice of medicine by developing methods and techniques for generating systematic reviews and clinical practice guidelines
- The main challenge for EBM remains how to develop a coherent theory of decision making by relating to other decision science disciplines

Search strategy and selection criteria

We searched PubMed for English language articles using the following search criteria: ("Evidence-Based Medicine/ethics"[Majr] OR "Evidence-Based Medicine/history"[Majr] OR "Evidence-Based Medicine/methods"[Majr] OR "Evidence-Based Medicine/standards"[Majr] OR "Evidence-Based Medicine/statistics and numerical data"[Majr] OR "Evidence-Based Medicine/trends"[Majr]). "Evidence-Based Medicine"[Mesh] AND (critical[All Fields] AND appraisal[All Fields]). The last search was performed on April 19, 2016 (there was no restriction on the start date). 6009 hits were identified. We supplemented the search by searching our personal libraries and the references of selected articles. The reviewers have also provided useful references. We selected articles of relevance for the main sections of the paper using our own judgment.

understanding the results of clinical studies, and determining how best to apply the results to their everyday practice.[9] Subsequently, detailed guidance published in journal articles and associated textbooks,[10,11] complemented by popular tools such as the Graphic Appraisal Tool for Epidemiology,[12] resulted in EBM becoming increasingly integrated into medical curricula worldwide.[13] Additionally, ratings of important developments in medicine have placed EBM on par with antibiotics and anaesthesia for the practice of medicine.[14] Here, we briefly review the philosophical underpinnings of EBM and, in more detail, its progress during the past quarter century. The current discussion goes beyond previous reviews,[15] placing the development of EBM within a framework of its historical and philosophical underpinnings; highlighting the role of EBM in the development of standards for clinical research and measuring practice; clearly documenting the important changes in EBM that occurred over more than two decades; addressing the critiques and limitations of EBM; and predicting the development of EBM in the next 25 years.

EBM and the theory of knowledge

On the surface, EBM proposes a specific association between medical evidence, theory, and practice. EBM does not, however, offer a new scientific theory of medical knowledge,[16,17] but instead has progressed as a coherent heuristic structure for optimising the practice of medicine, which explicitly and conscientiously[18] attends to the nature of medical evidence. Central to the epistemology of EBM is that what is justifiable or reasonable to believe depends on the trustworthiness of the evidence, and the extent to which we believe that evidence is determined by credible processes.[17] Although EBM acknowledges a role for all empirical observations, it contends that controlled clinical observations provide more trustworthy evidence than do uncontrolled observations, biological experiments, or individual clinician's experiences.

The basis for the first EBM epistemological principle is that not all evidence is created equal, and that the practice of medicine should be based on the best available evidence. The second principle endorses the philosophical view that the pursuit of truth is best accomplished by evaluating the totality of the evidence, and not selecting evidence that favours a particular claim.[16]

Evidence is, however, necessary but not sufficient for effective decision making, which has to address the consequences of importance to the decision maker within the given environment and context.[17] Thus, the third epistemological principle of EBM is that clinical decision making requires consideration of patients' values and preferences.

EBM's initial hierarchy of evidence

EBM originally focused on documenting biases in research applied to clinical practice, understanding the results of clinical studies, and considering situations (related to patient characteristics, family, and social and economic environment) in which these results can and cannot be usefully applied. In doing so, EBM addressed the need to identify poor research practices in how research is conceived, conducted, published, and used.

Several investigators have provided examples of biased research leading to suboptimal medical practice, lamenting the "scandal of poor medical research"[19] and claiming that "most research finding[s] are false".[20] Estimates suggest that 50% of research effort is wasted at each stage of generation and reporting of research, resulting in more than 85% of total research wasted;[21] the human toll of spurious research findings has equally been enormous.[22] For example, thousands of women underwent gruelling and sometimes fatal bone marrow transplantation for treatment of breast cancer based on biased research.[23] Promotion of prophylactic antiarrhythmic therapy in patients with myocardial infarction, based on physiological reasoning that suppression of arrhythmias would reduce mortality, proved disastrous—more Americans died from the use of these drugs than in the Vietnam war.[24] Millions of healthy women were prescribed hormone replacement therapy on the basis of hypothesised reduction in cardiovascular risk; randomised trials refuted these benefits and demonstrated that hormone replacement therapy increased the incidence of breast cancer.[25]

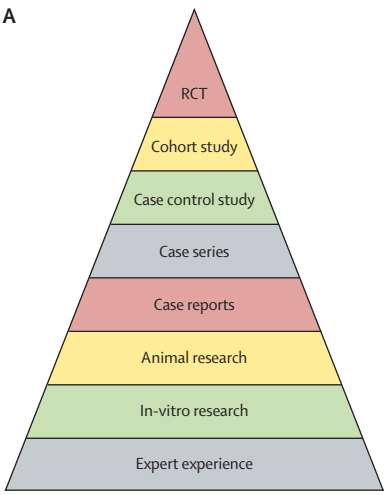
In response, EBM, from its inception, developed schemas for the assessment of the quality of evidence, reflecting the first EBM epistemological principle: the higher the quality of evidence, the closer to the truth are estimates of diagnostic test properties, prognosis, and the effects of health interventions.

Further, EBM writings acknowledged the challenges of understanding the quantitative results of clinical research, and of applying these results to patients who do not necessarily fit the eligibility criteria of the available studies. This work, focused on educating front-line clinicians, was so quickly acknowledged that, within a decade of their introduction, EBM principles became part of the core requirements for most undergraduate and postgraduate medical education worldwide.

The initial hierarchies of evidence that EBM proposed focused on the design of clinical studies and were relatively simple (figure 1A). For therapy, the hierarchy provided no equivocation regarding the superiority of randomised controlled trials (RCTs) over observational studies for determining the trustworthiness of evidence related to treatment effects—although early work fully acknowledged the limitations of small sample size and the questionable application of clinical findings, often based on surrogate markers, to patients who differed from those included in the primary studies.

Almost immediately, observers objected, noting that RCTs can also be biased, and hence should not automatically be equated with high-quality evidence.[28] As

A



B

Quality of evidence	Study design	Lower quality if*	Higher quality if†
High	Randomised trial	Study limitations - 1 serious - 2 very serious Inconsistency - 1 serious - 2 very serious Indirectness - 1 serious - 2 very serious Imprecision - 1 serious - 2 very serious Publication bias - 1 likely - 2 very likely	Large effect + 1 large + 2 very large Dose response + 1 evidence of a gradient All plausible confounders + Would reduce a demonstrated effect or + Would suggest a spurious effect when results show no effect
Moderate			
Low	Observational study		
Very low			

Figure 1: **Hierarchy of evidence: traditional EBM versus GRADE**
Comparison of traditional EBM hierarchy of evidence (1991–2004)[26] with GRADE classification of the quality of evidence (confidence, certainty; 2004 to present).[27] (A) Traditional EBM hierarchy of evidence. (B) GRADE classification of the quality of evidence. EBM=evidence-based medicine. GRADE=Grading of Recommendations Assessment, Development, and Evaluation. RCT=randomised controlled trial. *Quality of study moves down one or two grades. †Quality of study moves up one or two grades.

a result, during the first decade of the EBM movement, many authors published modifications of the original evidence hierarchy: by 2002, 106 systems to rate the quality of medical research evidence were available.[29] When investigators applied some of these quality instruments to a set of studies, the result was extensive disagreement, with ratings ranging, for the same studies, from excellent to poor.[30] One evaluation of these systems concluded that none was particularly useful for research or the practice of medicine, that their continued use will not reduce errors in making recommendations or improve communication between guideline developers and guideline users, and thus they will not help people make well informed decisions.[31]

Promoting the principle of totality of evidence—the rise of systematic reviews

Initial formulations of the hierarchy of evidence were also limited in that they confused the method of collecting evidence with the underlying study design. The view that "science is cumulative, and scientists should cumulate scientifically"[32] reflects the second EBM principle: health claims should be based on systematic reviews that summarise the best available evidence.[33] In keeping with this view, earlier formulations of the hierarchy placed systematic reviews at the top, followed by RCTs. This classification is misguided in that systematic reviews are a way of summarising the evidence, whereas RCTs are a type of study design. The distinction is vivid when one considers that systematic reviews can summarise not only RCTs, but also cohort studies, case-control studies, and even case reports.

The Cochrane Collaboration[34] represents the watershed movement responsible for the biggest advances in systematic review methodology. Named after Archie Cochrane, a visionary who demanded that the medical profession organise a "critical summary, by speciality or subspecialty, adapted periodically, of all relevant randomised controlled trials",[6] the Cochrane Collaboration has marshalled over 37 000 collaborators from more than 130 countries devoted to conducting systematic reviews.[35] Although reviews of RCTs remain the Collaboration's primary focus, their scope now includes observational studies addressing intervention effects, as well as diagnostic tests and prognostic models.

Iain Chalmers, the individual most responsible for the creation of the Cochrane Collaboration, has noted the "scandalous failure of science to cumulate evidence systematically"[36] and documented instances in which people have suffered and died unnecessarily, and resources for health care and health research have been wasted, because existing research evidence was not reviewed systematically.[32] When systematic reviews have been applied in a timely manner, they have resulted in major changes in the practice of medicine, including establishing standards of care for chemotherapy and hormonal therapy for early-stage breast cancer;[37,38] helping to overturn decades-old erroneous advice that infants should not sleep on their backs (and thus preventing sudden infant deaths);[39] and most recently shifting management of one of the world's most common disorders, community-acquired pneumonia, toward use of a short course of oral steroids.[40]

Systematic reviews, the most cited type of clinical research articles,[41] are essential for developing clinical practice guidelines, for avoiding duplication of research

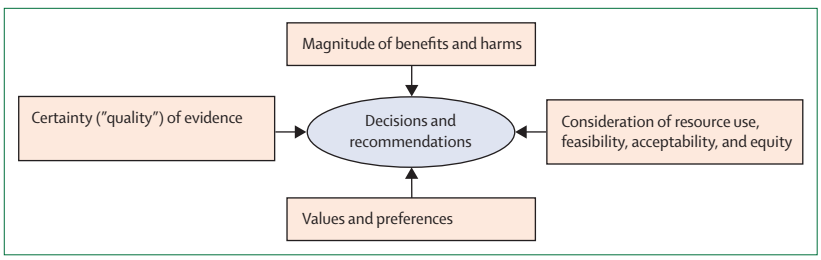


Figure 2: **Factors affecting decision making according to GRADE[27,56]**
GRADE=Grading of Recommendations Assessment, Development, and Evaluation.

efforts, and for helping inform design of new research studies. *The Lancet* has acknowledged the necessity for systematic summaries to inform new findings, demanding that authors of primary studies explain "the relation between existing and new evidence by direct reference to an existing systematic review or meta-analysis".[42] Further developments include an International Prospective Register of Systematic Reviews,[43] increasing sophistication of the methods,[44] and the increasing incorporation of systematic reviews into evidence tables and decision aids[45] to facilitate clinical decision making.

More sophisticated hierarchies and their application to clinical practice guidelines

As the awareness of the limitations of the initial simple hierarchies of evidence grew, another stream in the progress of EBM was occurring. A decade of efforts to teach EBM to medical trainees had revealed that few clinicians would ever have the skills—and those with the skills would seldom have time—to conduct sophisticated assessment of the evidentiary basis for their practice.[46] This realisation led to a refocusing of EBM efforts, directing clinicians to processed sources of evidence, and aiding decision making by advancing the science of trustworthy clinical practice guidelines that would be available to clinicians at the point of care delivery.

The focus on pre-processed evidence, and on clinical practice guidelines in particular, had other drivers. Initially neglected in EBM writings—with their focus on educating clinicians to read primary research studies—seminal efforts to put practice guidelines on a scientific footing had begun in the 1980s.[47,48] Subsequently, in 1990, recognition of unwarranted variation in medical practice[49] prompted the US Institute of Medicine (IOM) to call for standardisation of clinical practice via the development and application of clinical practice guidelines.[50] Problematic quality of care continues: estimates from the US suggest that more than 30% of health care is inappropriate or wasteful; between 70 000 and a third of all deaths[51] occur annually as a result of medical errors; and that only 55% of needed health services are delivered.[52]

Guidelines represent one strategy to address these problems: if guidelines are trustworthy—eg, according to IOM criteria,[53] which include a systematic review of evidence, explicit consideration of values and preferences, and addressing of issues related to conflicts of interest—adherence to them could prevent as many as a third of the leading causes of death, and reduce health-care spending by a third.[54]

These three realisations—the limitations of existing evidence hierarchies, the importance of processed evidence for ensuring evidence-based practice, and the related potential for practice guidelines to improve practice and outcomes—led to the development of a new approach to rating evidence quality and the grading strength of recommendations, termed the Grades of Recommendation Assessment, Development, and Evaluation (GRADE) system, which was first published in 2004.[27] The new system has enjoyed success similar to EBM itself: GRADE has been adopted by over 100 organisations, including the Cochrane Collaboration, the National Institute for Health and Care Excellence, WHO, and UpToDate.[27]

GRADE provides a much more sophisticated hierarchy of evidence (figure 1B), which addresses all elements related to the credibility of bodies of evidence: study design, risk of bias (study strengths and limitations), precision, consistency (variability in results between studies), directness (applicability), publication bias, magnitude of effect, and dose-response gradients. In doing so, GRADE protects against both superficial assessment and unwarranted confidence in RCTs, as well as dogmatic decisions. Further, the rapidly increasing use of GRADE has resulted, and will increasingly result, in marked improvement in the quality of systematic reviews.

GRADE allows not only for limitations in bodies of evidence from RCTs, but also the rating of observational studies as high-quality evidence (as in cases of dialysis, insulin for diabetic ketoacidosis, and hip replacements, for which RCTs have—appropriately—never been undertaken; figure 1B). GRADE, therefore, recognises the potential for observational studies to provide definitive causal evidence, particularly relevant for harmful exposures (eg, establishing that smoking causes lung cancer).

GRADE now provides guidance for assessing the quality of evidence, not only for management issues but also for diagnostic and prognostic issues, as well as animal studies and network meta-analyses.[55] GRADE has also addressed the process of moving from evidence to recommendations, beginning with summary of findings tables that present not only the quality of evidence, but estimates of both the relative and absolute effects for each patient-important outcome.[56,57] Core issues in that process include the magnitude of benefits, burdens and harms, quality of evidence (certainty or confidence in evidence), and values and preferences (relative importance of outcomes). Additional issues that guideline panels might consider include resource use (costs), feasibility, acceptability, and health equity (figure 2).

By presenting information in different formats, GRADE[58] and similar EBM initiatives[59,60] address framing effects (referring to the phenomenon of people making different decisions when identical information is

presented differently, in terms of gains *vs* losses). Through explicit consideration of judgment regarding the balance of benefits and harms, resource use, and issues of feasibility and equity, EBM has articulated a framework for rational decision making.[56,57,61]

The recognition that values and preferences vary widely among individuals has an important implication: the standardisation of care, which was one of the original reasons for the introduction of guidelines, and is still considered a key rationale for assessing the quality of care initiatives,[62] is neither possible nor desirable[63] for the many value and preference-sensitive decisions that clinicians and patients face. GRADE acknowledges the intrinsic variability of patient preferences in its classification of recommendations as strong (right for all, or almost all) and weak (conditional, contingent—right for most but not all, and requiring presentation of evidence that facilitates shared decision making).

Development of standards for conduct and reporting of clinical research

Scientific communities have embraced EBM-related initiatives to develop guidance and checklists for improving design, conduct, and reporting of research. Numerous such initiatives have occurred during the past 25 years, including checklists and statements on how to develop a research protocol and report randomised trials, observational studies, diagnostic test studies, predictive models, and genetic testing studies; they can be accessed via the EQUATOR website. Along the way, researchers have increasingly differentiated between explanatory (also known as mechanistic or proof-of-concept efficacy) trials that address the question "can intervention work in the ideal setting?" versus pragmatic (also known as practical, effectiveness) trials that address the question "does it work in real-world settings?" and "is it worth it and should it be paid for? (efficiency)".[64,65]

There is some evidence that these initiatives have resulted in improvement of the quality of reporting of research—for instance, the reporting of RCTs has improved as a result of the CONSORT checklist.[66] Optimal reporting is desirable, but worse than poor reporting is failure to report or suppression of clinical research. Currently, investigators report only 50% of their trials,[67,68] a major and avoidable threat to the body of scientific knowledge. When half of studies are unreported, both patient care and new research initiatives will often be flawed. Despite a longstanding awareness of the problem of publication bias,[69] the only possible solution to the problem—registration of all trial protocols before research is actually undertaken, and full reporting of the results in a timely manner after the study is completed—received only haphazard adherence in 2016.[70]

Evidence dissemination and access

Practising EBM at what David Sackett called "the coal-face"[71] of clinical care requires rapid access to the best available evidence, suitably filtered to ensure efficient use. Provision of that access is challenging and has been one of the most important academic endeavours of EBM. The process is complicated by the ongoing information explosion: estimates in 2000 suggested that more than 6 million articles are published in more than 20 000 biomedical journals every year.[72] MEDLINE alone contains over 22 million indexed citations from more than 5600 journals,[73] and 75 RCTs and 11 systematic reviews are published every day.[74]

Dealing with both the information explosion and inherent human brain limitations in processing evidence has required the application of the EBM principle of critical appraisal to identify high-quality research on the clinicians' behalf. Haynes and colleagues[75] have developed a model service that uses EBM critical appraisal techniques to systematically evaluate more than 3000 articles per year from all medical disciplines. Using EBM information processing and filtering, they reported that, on average, clinicians need to be aware of only about 20 new articles per year (99·96% noise reduction) to keep up to date, and, to stay up to date in their area of expertise, authors of evidence-based topics need be aware of only five to 50 new articles per year.[75] Similarly, practising oncologists need be cognisant of only 1–2% of published evidence that is valid and relevant to their practice.[76]

Information services provide clinicians with alerts when this key new information appears in the medical literature, as well as providing filtered search systems that prioritise processed evidence (including clinical practice guidelines and systematic reviews). Electronic textbooks also provide valuable pre-processed information, including evidence-based summaries and GRADE recommendations (eg, Dynamed and UpToDate), as do other evidence summaries (eg, Best Evidence in Emergency Medicine).[45] However, an evaluation of 26 existing point-of-care information summaries found uneven quality across the products, with some products scoring higher on evidence-based dimensions than others.[77]

Nevertheless, electronic platforms based on the GRADE framework (eg, Making GRADE the Irresistible Choice[78,79]) that allow digitally structured storage of information are likely to play an important role in facilitating the creation, dissemination, and dynamic updating of trustworthy evidence summaries, guidelines, and decision aids. Such platforms also facilitate rapid updating of systematic reviews and guideline adaptation,[79] automated publication of multi-layered presentation formats on smart phones and other devices, and integration of evidence (summaries) and recommendations into electronic medical records as decision support systems. Formal research has optimised and will continue to improve presentation formats of evidence summaries and recommendations in these applications, helping to ensure maximal uptake by front-line clinicians.

For more on **EQUATOR** see http://www.equator-network.org/index.aspx?o=1032

Development of tools to improve decision making

The many factors that determine peoples' decision making can be classified as the effect of (1) context or framing, (2) situational or contextual factors (eg, psychosocial context or characteristics of the health-care system), and (3) individual characteristics of a decision maker (eg, experience, cultural background, and values and preferences).[80,81] The individual characteristics of a decision maker relates to the third principle of EBM: evidence never determines decisions; it is always evidence in the context of values and preferences.

The third principle of EBM is in keeping with a cultural change in medicine over the past 20 years: the growing emphasis on patient autonomy, and the associated priority given to shared decision making. Although widely acknowledged as desirable, the challenges to the implementation of shared decision making remain formidable. Health-care providers face severe time constraints and might not have the relevant evidence readily available or the skills necessary to optimally engage patients.

Decision aids that communicate harms, benefits, and alternatives in an easily understood manner represent a possible solution to the challenges of shared decision making.[82] These too face challenges: they are often based on inadequate and inaccurate evidence summaries from the start; if optimally evidence-based at the start, they fail to update appropriately; and, designed essentially as patient information, they often achieve little in the way of facilitating useful discussion between clinicians and patients.[83]

Point-of-care decision aids specifically designed for the clinician—patient encounter show promise for advancing the shared decision making cause. When created from the previously mentioned electronic platforms, developers can access and present the best updated evidence for the clinician to share on electronic devices.[45] Formal user testing has provided a format that allows the developer to address the two other determinants of decision making introduced at the beginning of this section: framing of the information and ensuring relevance to the particular clinical setting. Further development, testing, and dissemination of point-of-care decisions aids represents a frontier for future EBM advances.

Criticism of EBM

Persistent criticisms of EBM have focused on three major issues. The first argues that EBM relies on reductionism of the scientific method;[84,85] critics have been particularly vocal regarding overly strict adherence to the evidence hierarchy pyramid (figure 1A), which they viewed as narrow and simplistic.[28,84–86] It took almost 15 years for EBM to respond fully to this legitimate concern; the sophisticated hierarchy of evidence offered by the GRADE framework effectively addresses the issue (figure 1B).

The second claim is that EBM encourages formulaic "cookbook medicine",[87] discouraging deliberation and clinical reasoning and leading to automatic decision making. This criticism was reframed in a recent article that raised the question of whether EBM is a "movement in crisis", and issued warnings regarding approaches that are excessively algorithmic (in the process, perhaps neglecting the frequent usefulness of algorithms).[88] The critics have noted that care for a particular patient "may not match what the best (average) evidence seems to suggest."[88] These[88] and other authors[89] lament that EBM has neglected the humanistic and personal aspects of medical care and moved the focus away from the individual.[90] In reality, EBM has aggressively promoted the need to consider a patient's values in every preference-sensitive decision.[91] A focus on individual patient values, which involve how patients view the world and their relationships with their environment, friends, and loved ones, lies at the heart of the humanistic practice of medicine.

Notably, from its early days, EBM has focused on the individual patient. Aspects of that focus included championing randomised trials in individual patients (N-of-1 randomised trials),[92] highlighting differences in baseline risk (large effects in patients with high baseline risk and small effects in those with low baseline risk), and providing guides for the credibility of subgroup analysis.[93]

A third criticism is that EBM promotes rule-based reasoning instead of intuitive and experiential thinking, which characterise expert judgment.[88] EBM has indeed maintained that scientific evidence should reflect knowledge that is publicly shared and easily understood by all qualified professionals in the field.[16] It is understandable that EBM's stress on the use of results of replicable research could be interpreted as diminishing the role of expertise and judicious clinical judgment. EBM does, in fact, highly value the critical role of expertise in health-care-delivery by emphasising the importance of judicious judgment in critical appraisal and decision making.

Another criticism of EBM is that there is no high-quality evidence that its application has improved patient care. We would rebut by noting the history of a decade-or-more delays in implementing interventions, such as thrombolytic therapy for myocardial infarction,[94] and highlighting the previously described examples of routinely administered useless and harmful interventions, such as lidocaine to patients after myocardial infarction, placing infants on their stomachs to sleep, or hormone replacement therapy for postmenopausal women, that preceded the widespread implementation of EBM.

Recent writings have also claimed that EBM has been "high-jacked"[95] by commercial interests that, having learned how to exploit EBM principles, have been creating doubt when none reasonably exists,[96] spinning the message,[97] and medicalising issues that are better

viewed as the natural accompaniment of the human experience. It is true, for instance, that most compelling randomised trials highlighted in EBM-oriented texts have been conducted by the pharmaceutical industry. EBM writings have provided guides for detecting misleading study designs and interpretation—for example, choosing an inferior comparator[98] or undertaking mega-trials and then misrepresenting very small effects as major breakthroughs.[99] The extent to which such warnings and guides have adequately protected clinicians from misleading presentations is at best questionable, and probably limited. EBM practitioners in particular, and the medical community in general, need to continue to push back against these distortions that often result in "too much medicine".[100]

However, no critic ever suggested that reliable evidence should not be a key to effective problem solving and decision making. Humans are "informavores"[101]—we need evidence to effectively function in the world around us.

The next 25 years

EBM has disseminated three major tenets: an increasingly sophisticated hierarchy of evidence, the need for systematic summaries of the best evidence to guide care, and the requirement for considering patient values in important clinical decisions. EBM has contributed to, and perhaps spawned, a number of related initiatives. These initiatives include a focus on comparative effectiveness research,[102] over or under diagnosis and over or under treatment,[100] measurements of the quality of care,[103] improving publishing standards,[104] ensuring all trials are registered,[70,105] and avoiding waste in research production, including discontinued use of misguided interventions that have become part of established practice.[106] These initiatives reflect the broad scope of the EBM movement, which has expanded to include disciplines such as nursing, dentistry, public health, and health policy (so-called evidence-based health care), as well as recognition for the need of evidence-based implementation science to ensure optimal function of clinics, hospitals, and health systems.[52]

EBM will have to address several challenges in the next quarter century. Failure to publish, and indeed suppression of research results, remains a problem. Optimal delivery of clinical care requires far more efficient production and rapid dissemination of both systematic reviews and practice guidelines.[107] Achieving this goal will require further advances in building experienced research teams geared toward rapid turnaround in creating rigorous evidence summaries, automatised and text-mining software,[108] and the widespread adoption of electronic platforms that greatly facilitate rapid updating. Dissemination must include patient-friendly and clinician-friendly electronic access on all types of devices, including smart phones, to electronic medical records and, particularly for patients, social media.[45,109,110]

EBM will need to address the place of evidence generated by so-called big data[111] mining in relation to traditional observational studies and randomised trials for the development of a "continuously learning health care system".[112] EBM has yet to generate a coherent theory of health-care decision making[113] and will continue to partner with other disciplines, such as cognitive and decision sciences, toward this goal.[114] On a more practical level, major challenges remain in providing clinicians with tools to make shared decision making at the point of care entirely practical and efficient, and a positive experience for both patients and clinicians.

In conclusion, efforts are well underway in each of the problematic areas of EBM, and progress is certain. Whatever the extent of future progress, EBM's success in providing a framework for fully integrating research evidence into the delivery of health care,[115] and raising awareness of the need for consideration of individual patient values and preferences, will remain enduring contributions to clinical medicine and related fields.

Contributors

BD wrote the first draft, which was extensively revised by GHG. Both authors contributed equally to the concept of the paper, data interpretation, and writing. The final version of the paper has been agreed by both authors.

Declaration of interests

Both authors have devoted significant aspects of their careers to the development of evidence-based practice.

References

1 Lind J. A treatise of scurvy. In three parts. Containing an enquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. Edinburgh: Printed by Sands, Murray, and Cochran, 1753.
2 Matthews JR. Quantification and quest for medical certainty. Princeton, NY: Princeton University Press, 1995.
3 Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J* 1979; **121:** 1193–254.
4 Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1986; **89** (2 suppl)**:** 2S–3S.
5 Eddy DM, Billings J. The quality of medical evidence: implications for quality of care. *Health Aff (Millwood)* 1988; **7:** 19–32.
6 Cochrane A. Effectiveness and efficiency: random reflections on health services. London: Nuffield Provincial Hospitals Trust, 1972.
7 Eddy DM. Guidelines for the cancer-related checkup: recommendations and rationale. *CA Cancer J Clin* 1980; **30:** 193–240.
8 Guyatt G. Evidence-Based Medicine. *ACP J Club* 1991; **A-16:** 114.
9 Evidence-based medicine working group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992; **268:** 2420–25.
10 Straus S, Richardson W, Glasziou P, Haynes R. Evidence-based medicine. How to practice and teach EBM. Edinburgh: Churchill Livingstone, 2005.
11 Guyatt G, Drummond R, Meade MO, Cook DJ. Users' guides to the medical literature: a manual for evidence-based clinical practice. Boston, MA: McGraw-Hill, 2014.
12 Jackson R, Ameratunga S, Broad J, et al. The GATE frame: critical appraisal with pictures. *Evid Based Med* 2006; **11:** 35–38.
13 Maggio LA, Tannery NH, Chen HC, ten Cate O, O'Brien B. Evidence-based medicine training in undergraduate medical education: a review and critique of the literature published 2006–2011. *Acad Med* 2013; **88:** 1022–28.
14 Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. *BMJ* 2007; **334** (suppl 1)**:** s10–13.
15 Montori VM, Guyatt GH. Progress in evidence-based medicine. *JAMA* 2008; **300:** 1814–16.

16 Djulbegovic B, Guyatt GH, Ashcroft RE. Epistemologic inquiries in evidence-based medicine. *Cancer Control* 2009; **16:** 158–68.

17 Djulbegovic B, Guyatt G. EBM and the theory of knowledge. In: Guyatt G, Meade M, Cook D, eds. Users' guides to the medical literature: a manual for evidence-based clinical practice. Boston, MA: McGraw-Hill, 2014.

18 Sackett D, Rosenberg W, Muir Gray J, Haynes R, Richardson W. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; **312:** 71–72.

19 Altman D. The scandal of poor medical research. *BMJ* 1994; **308:** 283–84.

20 Ioannidis J. Why most published research findings are false. *PLoS Med* 2005; **2:** e124.

21 Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. *Lancet* 2014; **383:** 101–04.

22 Chalmers I, Bracken MB, Djulbegovic B, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014; **383:** 156–65.

23 Rettig RA, Jacobson PD, Farquhar CM, Aubry WM. False hope. Bone marrow transplantation for breast cancer. New York, NY: Oxford University Press, 2007.

24 Moore TJ. Deadly medicine. New York, NY: Simon & Schuster, 1995.

25 Rossouw J, Anderson G, Prentice R, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. *JAMA* 2002; **288:** 321–33.

26 Guyatt G, Sackett D, Sinclair J, et al. User's guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1995; **274:** 1800–04.

27 Atkins D, Best D, Briss P, et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004; **328:** 1490.

28 Worrall J. What evidence in evidence-based medicine? *Philos Sci* 2002; **69:** S316–30.

29 West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment No 47. Rockville, MD: Agency for Health Care Research and Quality, 2002.

30 Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; **282:** 1054–60.

31 Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. *BMC Health Serv Res* 2004; **4:** 38.

32 Chalmers I. The lethal consequences of failing to make use of all relevant evidence about the effects of medical treatments: the need for systematic reviews. In: Rothwell P, ed. Treating individuals: from randomised trials to personalised medicine. London: Lancet, 2007: 37–58.

33 Chalmers I. Addressing uncertainties about the effects of treatments offered to NHS patients: whose responsibility? *J R Soc Med* 2007; **100:** 440–41.

34 Chalmers I. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993; **703:** 156–63.

35 Cochrane Collaboration. About Us. 2016. http://www.cochrane.org/about-us (accessed Jan 3, 2017).

36 Chalmers I. The scandalous failure of scientists to cumulate scientifically. *Clin Trials* 2005; **2:** 229–31.

37 Early Breast Cancer Trialists' Collaborative Group. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *Lancet* 2012; **379:** 432–44.

38 Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015; **386:** 1341–52.

39 Gilbert R, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. *Int J Epidemiol* 2005; **34:** 874–87.

40 Siemieniuk RA, Meade MO, Alonso-Coello P, et al. Corticosteroid therapy for patients hospitalized with community-acquired pneumonia: a systematic review and meta-analysis. *Ann Intern Med* 2015; **163:** 519–28.

41 Patsopoulos NA, Analatos AA, Ioannidis JPA. Relative citation impact of various study designs in the health sciences. *JAMA* 2005; **293:** 2362–66.

42 Young C, Horton R. Putting clinical trials into context. *Lancet* 2005; **366:** 107–08.

43 PROSPERO. International Registry of Systematic Reviews. http://www.crd.york.ac.uk/prospero (accessed April 21, 2016).

44 Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. Finding what works in health care: standards for systematic reviews. Washington, DC: The National Academies Press, 2011.

45 Agoritsas T, Heen A, Brandt L, Alonso-Coello P, Kristiansen A, Akl E. Decision aids that really promote shared decision making: the pace quickens. *BMJ* 2015; **350:** g7624.

46 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:** 954–55.

47 Eddy DM. Designing a practice policy. Standards, guidelines, and options. *JAMA* 1990; **263:** 3077–84.

48 Eddy DM. Practice policies: guidelines for methods. *JAMA* 1990; **263:** 1839–41.

49 Wenneberg J. Which rate is right? *N Engl J Med* 1986; **314:** 310–11.

50 Institute of Medicine (US) Committee to Advise the Public Health Service on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new agency. Field MJ, Lohr KN, eds. Washington, DC: National Academy Press, 1990.

51 Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *BMJ* 2016; **353:** i2139.

52 Djulbegovic B. A framework to bridge the gaps between evidence-based medicine, health outcomes, and improvement and implementation science. *J Oncol Pract* 2014; **10:** 200–02.

53 Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Clinical practice guidelines we can trust. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, eds. Washington, DC: National Academies Press, 2011.

54 Pronovost PJ. Enhancing physicians' use of clinical guidelines. *JAMA* 2013; **310:** 2501–02.

55 GRADE Working Group. List of GRADE working group publications and grants. 2016. http://www.gradeworkinggroup.org (accessed Aug 21, 2016).

56 Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. *BMJ* 2008; **336:** 1049–51.

57 Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines 15: going from evidence to recommendation—determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013; **66:** 726–35.

58 Guyatt GH, Oxman AD, Santesso N, et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. *J Clin Epidemiol* 2013; **66:** 158–72.

59 Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making* 2007; **27:** 696–713.

60 Schwartz L, Woloshin S, Welch H. Using a drug facts box to communicate drug benefits and harms. *Ann Intern Med* 2009; **150:** 516–27.

61 Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336:** 924–26.

62 IOM. Vital signs: core metrics for health and health care progress. Washington, DC: The National Academies Press, 2015.

63 Djulbegovic B, Guyatt GH. Evidence-based practice is not synonymous with delivery of uniform health care. *JAMA* 2014; **312:** 1293–94.

64 Porzsolt F, Rocha NG, Toledo-Arruda AC, et al. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages. *Pragmat Obs Res* 2015; **6:** 47–54.

65 Djulbegovic B, Paul A. From efficacy to effectiveness in the face of uncertainty indication creep and prevention creep. *JAMA* 2011; **305:** 2005–06.

66 Moher D, Jones A, Lepage L, CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials. A comparative before-and-after evaluation. *JAMA* 2001; **285:** 1992–95.

67 Chalmers I, Glasziou P, Godlee F. All trials must be registered and the results published. *BMJ* 2013; **346:** f105.

68 McGauran N, Wieseler B, Kreis J, Schuler YB, Kolsch H, Kaiser T. Reporting bias in medical research—a narrative review. *Trials* 2010; **11:** 37.

69 Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986; **4:** 1529–41.

70 Goldacre B. Make journals report clinical trials properly. *Nature* 2016; **530:** 7.

71 Sackett DL. Why randomized controlled trials fail but needn't: 1. failure to gain "coal-face" commitment and to use the uncertainty principle. *CMAJ* 2000; **162:** 1311–14.

72 Djulbegovic B, Lyman GH, Ruckdeschel J. Why evidence-based oncology? *Evidence-based Oncology* 2000; **1:** 2–5.

73 MEDLINE Fact Sheet. 2016. https://www.nlm.nih.gov/pubs/factsheets/medline.html (accessed April 20, 2016).

74 Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010; **7:** e1000326.

75 Haynes RB, Cotoi C, Holland J, et al. Second-order peer review of the medical literature for clinical practitioners. *JAMA* 2006; **295:** 1801–08.

76 Vincent S, Djulbegovic B. Oncology treatment recommendations can be supported only by 1-2% of published high-quality evidence. *Cancer Treat Rev* 2005; **314:** 319–22.

77 Kwag KH, González-Lorenzo M, Banzi R, Bonovas S, Moja L. Providing doctors with high-quality information: an updated evaluation of web-based point-of-care information summaries. *J Med Internet Res* 2016; **18:** e15.

78 Vandvik P, Brandt L, Alonso-Coello P, Treweek S, Akl E, Kristiansen A. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. *Chest* 2013; **144:** 381–89.

79 Kristiansen A, Brandt L, Agoritsas T, et al. Applying new strategies for the national adaptation, updating, and dissemination of trustworthy guidelines: results from the Norwegian adaptation of the Antithrombotic Therapy and the Prevention of Thrombosis, 9th edn: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2014; **146**: 735–61.

80 Appelt KC, Milch KF, Handgraaf MJJ, Weber EU. The decision making individual differences inventory and guidelines for the study of individual differences in judgment and decision-making research. *Judgm Decis Mak* 2011; **6:** 252–62.

81 Hastie R, Dawes RM. Rational choice in an uncertain world. 2nd edn. Los Angeles, CA: Sage Publications, Inc., 2010.

82 A to Z Inventory of Decision Aids. https://decisionaid.ohri.ca/azinvent.php. (accessed Aug 21, 2016).

83 Hargraves I, LeBlanc A, Shah ND, Montori VM. Shared decision making: the need for patient-clinician conversation, not just information. *Health Aff (Millwood)* 2016; **35:** 627–29.

84 Miles AM, Bentley P, Polychronis A, Grey J, Price N. Advancing the evidence-based healthcare debate. *J Eval Clin Pract* 1999; **5:** 97–101.

85 Miles A, Bentey P, Polychronis A, Grey J. Evidence-based medicine: why all the fuss? This is why. *J Eval Clin Pract* 1997; **3:** 83–86.

86 Miles A, Grey J. New perspectives in the evidence-based healthcare debate. *J Eval Clin Pract* 2000; **6:** 77–84.

87 Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. *Health Aff (Millwood)* 2005; **24:** 18–28.

88 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? 2014; **348:** g3725.

89 Miles A, Asbridge JE, Caballero F. Towards a person-centered medical education: challenges and imperatives. *Educ Med* 2015; **16:** 25–33.

90 McCartney M, Treadwell J, Maskrey N, Lehman R. Making evidence based medicine work for individual patients. *BMJ* 2016; **353:** i2452.

91 Guyatt GH, Haynes RB, Jaeschke RZ, et al. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. *JAMA* 2000; **284:** 1290–96.

92 Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. *Ann Intern Med* 1990; **112:** 293–99.

93 Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010; **340:** c117.

94 Lau J, Antman EM, Jimenez-Silva J, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med*; **327:** 248–54.

95 Ioannidis JPA. Evidence-based medicine has been hijacked: a report to David Sackett. *J Clin Epidemiol* 2016; **73:** 82–86.

96 Michaels D. Doubt is their product. *Sci Am* 2005; **292:** 96–101.

97 Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. *PLoS Med* 2008; **5:** e230.

98 Mann H, Djulbegovic B. Comparator bias: why comparisons must address genuine uncertainties. *J R Soc Med* 2013; **106:** 30–33.

99 Carrasco-Labra A, Montori VM, Ioannidis JPA, et al. Misleading presentations of clinical trial results. In: Guyatt GRD, Meade M, Cook D, eds. Users' guides to the medical literature: a manual for evidence-based clinical practice. 3rd edn. New York, NY: McGraw-Hill, 2014.

100 Moynihan R, Henry D, Moons KGM. Using evidence to combat overdiagnosis and overtreatment: evaluating treatments, tests, and disease definitions in the time of too much. *PLoS Med* 2014; **11:** e1001655.

101 Pirolli P, Card S. Information foraging. *Psychol Rev* 1999; **106:** 643–75.

102 Lohr KN. Comparative effectiveness research methods: symposium overview and summary. *Med Care* 2010; **48** (6 suppl): S3–S6.

103 Berenson RA, Pronovost PJ, Krumholz HM. Achieving the potential of health care performance measures. 2013. http://www.nejm.org/doi/full/10.1056/NEJMp1312287 (accessed Aug 21, 2016).

104 Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med* 2010; **8:** 24.

105 Antes G, Chalmers I. Under-reporting of clinical trials is unethical. *Lancet* 2003; **361:** 978–79.

106 Elshau AG, Watt AM, Mundy L, Willis CD. Over 150 potentially low-value health care practices: an Australian study. *Med J Aust* 2012; **197:** 556–60.

107 Schünemann HJ, Moja L. Reviews: Rapid! Rapid! Rapid!...and systematic. *Syst Rev* 2015; **4:** 1–3.

108 Paynter RA, Bañez LL, Berliner E, et al. EPC methods: an exploration of the use of text-mining software in systematic reviews. Research White Paper. Rockville, MD: Agency for Healthcare Research and Quality (US), 2016.

109 Guyatt G, Vandvik PO. Creating clinical practice guidelines: problems and solutions. *Chest* 2013; **144:** 365–67.

110 Vandvik PO, Brandt L, Alonso-Coello P, et al. Creating clinical practice guidelines we can trust, use, and share: a new era is imminent. *Chest* 2013; **144:** 381–89.

111 Sim I. Two ways of knowing: big data and evidence-based medicine. *Ann Intern Med* 2016; **164:** 562–63.

112 IOM (Institute of Medicine). Best care at lower cost: The path to continuously learning health care in America. Washington, DC: The National Academies Press, 2013.

113 Eddy DM. Evidence-based medicine: a unified approach. *Health Aff (Millwood)* 2005; **24:** 9–17.

114 Hozo I, Djulbegovic B, Luan S, Tsalatsanis A, Gigerenzer G. Towards theory integration: threshold model as a link between signal detection theory, fast-and-frugal trees and evidence accumulation theory. *J Eval Clin Pract* 2015; published online Dec 18. DOI:10.1111/jep.12490.

115 Karanicolas PJ, Kunz R, Guyatt GH. Point: evidence-based medicine has a sound scientific base. *Chest* 2008; **133:** 1067–71.