Received: 13 August 2020 | Accepted: 14 August 2020

DOI: 10.1111/jep.13474

Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

WILEY

EDITORIAL

Relevance, validity, and evidential reasoning in clinical practice

The renowned Canadian philosopher of science and probability, Ian Hacking, once quipped that "evidence is one thing that points beyond itself."[1(p37)] Well, evidence frequently does point to a given theory, hypothesis, decision, and so forth, but it need not. What Hacking's pithy remark overlooks is that the most reliable piece of evidence may not be persuasive (in the sense of not pointing beyond itself), and, conversely, evidence that is unreliable on most any measure may play a pivotal role in decision-making. Although the distinction between the goodness (reliability/validity) of evidence and the relevance (weight, importance) of evidence is a staple in legal practice, this distinction is frequently obscured in the literature on clinical decision-making.* One can see such obscuring in the privileged position afforded randomized controlled trials (RCTs), considered the most reliable and valid evidence by the evidence-based medicine (EBM) movement. Effect sizes derived from RCTs, we are told, are considered applicable in clinical populations unless there are "compelling reason to think that the biology in the population of interest is so different from that of the population tested that the magnitude of effect will differ substantially"[3(p1304)]—but such compelling reasons are considered rare by proponents of EBM.† In this editorial, we look at how is the distinction between goodness (reliability/validity) and relevance (weight/importance) factors into the enormous body of literature on the pros and cons of EBM and the theory of evidence that undergirds the EBM movement. We will proceed on the assumption that evidence may give empirical support to a position (whether a hypothesis, a theory, or a clinical decision), but to do, so it will require a conceptual framework that includes (at a minimum) (a) criteria governing the goodness (reliability/validity) of evidence elicited in support of the position; and (b) criteria governing the relevance (importance/weight) of that evidence to the position.

EBM has been widely (and rightly) criticized for its rigid and narrow conceptualization of evidence, but the fact is that it satisfies both the requirements of goodness and relevance, though in a manner that has struck many commentators as heavy-handed. Indeed, the guiding assumption of EBM is that it "requires the integration of the best research evidence with our clinical expertise and our patient's unique values and circumstances."[4(p1)] The characterization of "best research evidence" is itself piggybacked on a theory of evidence that is succinctly expressed in the fourth edition of *Evidence-based medicine: How to Practice and Teach It* as follows: "by best research evidence we mean valid [reliable] and clinically relevant research." The evidential power accorded to "clinically relevant research" has now been reformulated as The First Principle of EBM: "our inferences (and decisions) are best informed by systematic reviews (ie, syntheses of the totality of relevant high quality studies)."[5(p16)]

Such statements, presented in the previous paragraph, give us a glimpse of the theory of evidence, which is nowhere clearly articulated in the copious body of EBM literature, but is the motor of EBM and the lightning rod for resistance to the EBM movement. For example, one distinctive feature of the statement by Strauss et al. is that it bundles together two very thorny concepts (goodness and relevance) that are firmly set apart in other jurisdictions (eg, the law).‡ A desiderata of any theory of evidence is that it delimits what is to count as evidence (ie, sources of evidence). EBM takes this a step further by limiting the scope of these sources to the "best research evidence," along with a hierarchical ranking of the empirical support offered by evidence from these different sources according to their supposed validity, which is to enable the clinician to weigh and balance different pieces of evidence with the goal of improving clinical practice; that is, the ranking ensures that clinical decisions pass (a preconceived notion of) evidentiary muster. The caveat that the scope of evidence is limited to "clinically relevant research" effectively polices evidential reasoning, so that the body of evidence that in principal can be persuasive is restricted to discrete pieces of evidence that are valid/reliable according to the evidence hierarchies of EBM (see[6]).

Critics of EBM have been vocal in their opposition to the operationalizing of EBM, not merely the evidential authority that has been conferred on randomized control trials[7(p318)] but also to the stipulation that what is determined to be high quality on an EBM framework is the only evidence that is relevant in clinical decision-making (see[8-10]). These critics have repeatedly pressed the need for EBM to be open to other sources of medical knowledge—especially in light of acknowledgment by the EBM community that patient values, preferences, and circumstances (information that, for individual patients, is often not derived or even derivable from clinical trials) are important considerations and thus, relevant to clinical decision-making. With its restrictive conceptualization (and hierarchical ranking) of the sources of medical knowledge, some clinicians have voiced the opinion that EBM is an inappropriate response in many fields, notably, psychiatry, given "the character and complexity of psychiatric disorders and treatments."[8(p53)] In effect, what is being claimed here is that if the elements of the theory of evidence is being constructed with full regard

*A notable exception is Cartwright,[2] who (more than a decade ago) explicitly challenged philosophers of medicine to articulate practical advice about "how to determine when one empirical claim is relevant to the truth of another; that is, about empirical relevance" (p127).

†While there are no doubt other considerations within the EBM model for relevance, such statements can confuse the issue by seemingly conflating methodological rigour with relevance.

‡As will be argued in a later section, conceptualizing relevance in terms of goodness (validity/reliability) is the source of the many issues at play in the failure of EBM to curb scepticism about the promised integration of the patient "unique values and circumstances."

for the heterogeneity of clinical practice, what counts as the best clinical evidence will change from field to field, and, at the end of the day, the weight (relevance) EBM attaches to "syntheses of the totality of relevant high-quality studies" is just one point on a bell curve in clinical decision-making. It is possible to respect the diversity of evidence and reflect the heterogeneity of clinical practice if we set aside "evidence-based medicine" and instead conceptualize the relationship between evidence and clinical practice as "evidence informed medicine."[8,11§]

The COVID-19 pandemic has shone a beacon on the unwelcome consequences of the restricted view on evidence that persists in the EBM framework. Consider the debate about face coverings. Early on, the World Health Organization (WHO), a strong supporter of EBM, advocated that the public wearing of masks conferred no appreciable benefit in curbing the transmission of COVID-19—a claim that was nearly universally parroted by public health officials (including the Surgeon General of the United States), leading to the erosion of public's trust in public health (see[13]). There was (and still is) a dearth of RCT evidence examining the issue directly; the absence of evidence produced by RCTs was taken as no evidence that face coverings provide a social benefit. Extant trials were often compromised on EBM criteria of methodological rigor or were focused on other infectious agents (eg, severe acute respiratory syndrome [SARS], Middle East respiratory syndrome [MERS], influenza) and on source protection (rather than source control). Greenhalgh[14(p1071)] claims that "studies on the efficacy of masks in protecting the wearer are therefore irrelevant to the issues of source control." Greenhalgh has articulated how a wider view on evidence can benefit our understanding of the impact of face coverings (itself, a wider notion of "mask") in curbing COVID-19 transmission, while still falling within a rational purview of "evidence based." How such information falls into a category of "relevant evidence" within an EBM framework is not clear, although time may tell if the EBM community would be right to reject such reasoning. Still, one must wonder if strict adherence to an EBM framework provides a necessary safeguard or unnecessarily hamstrings public health officials in situations of imminent threat.

The lines of criticism we have so far presented are compelling and well-documented. However, there is another line that needs attention. Let us now turn our focus to the impact a theory of evidence that does not discriminate between goodness and relevance may have on the encounter with the individual patient. EBM has long acknowledged the importance of patient context (eg, values, preferences, circumstances) in clinical decision-making. Despite this acknowledgment, EBM, in our view, has not made sufficient headway on the promissory note of integrating the best science with patients "unique values and circumstances." This failure has been portrayed as a residue of paternalism in EBM and in other ways. This may very well be the case; EBM may be a paternalistic emperor in new clothing. However, if the focus is on EBM as a theory of evidence, it seems that this failure is the by-product of the lack of recognition that goodness and relevance are distinct criteria in a theory of evidence. Consider the case of the individual patient in the clinical encounter. Certainly, the patient's values and preferences are relevant to the decision, lest one ignore the basic principles of EBM advocated by its founders and continuously promoted in the literature. The question is whether the derivation of such information from the patient directly is indeed reliable and valid according to an EBM notion of evidence? The GRADE framework,[15] for example, advances a notion that information on patient values and preferences may be (or rather, is "ideally:) systematically derived from populations (ie, "typical" patients), and when coupled with strong evidence of therapeutic effectiveness, one might have grounds (where a "strong" recommendation is determined) to eschew "thorough (or even cursory) review of underlying evidence" and "detailed discussion with the patient,"[16(p50)]; also see.[17] The "stronger" evidence (on methodological grounds) from population studies thus becomes more relevant than that which could be derived from the patient directly. Again, a wider notion of evidence, that decouples relevance from goodness, may do much to ease such tensions as described here.

Although the EBM movement has elaborated a very specific (and admittedly controversial) framework for grading evidence for treatment outcomes in terms of its goodness (eg, see[18-20]), this movement has been comparatively silent on issues surrounding how theories about the goodness and relevance of evidence are to be integrated into clinical decision-making. Much has been written by proponents of EBM about what Worrall[7(p317)] characterized as "the evidential power of randomization," but little has been said about the evidential power of clinical judgment and expertise. What can be said is that however confident a clinician may feel about options for treatment, any recommendation made by the clinician not only reaches beyond the evidence (as Hacking notes) but also is shaped and ordered by the clinician's understanding of the treatment decision that stands the best chance of maximizing the patient's medical outcome. Following Veatch,[21] this understanding can be characterized as the patient's well-being.[¶] In an encounter (eg, the Emergency Department) where "doctor still knows best."[21(p703)] the relevance criterion might be satisfied by the clinician's belief that the evidence that best supports the treatment decision in fact maximizes the patient's medical outcome (however, this outcome is understood by the clinician).[#] It is the clinician's sense that a specific treatment will best maximize the patient's well-being that will effectively serve as the criterion that determines the relevance of evidence

[§]These criticisms are typically based on another way of conceptualizing evidence in terms of the totality of evidence, namely, as the sum of all considerations that tell in favour of a clinical decision. See, for example, Lipsitch,[12] who states that scientists "should keep their eyes open for any kind of information that can help them solve problems. Deciding, on principle, to reject some kinds of information outright, or to consider only particular kinds of studies, is counterproductive."

[¶]Veatch[21] takes great pains to highlight the tension between what he characterizes as the patient's "total well-being" and clinical well-being: "We now know that even in the ideal case physicians generally have no basis for knowing what would benefit their patients. Even if they can accurately diagnose disease and prognosticate its future course under various treatment options, they still cannot be expected to have any basis for knowing that one outcome is better than another for the patient who presents the medical problem to them."[21(p703)] Although Veatch does an admirable job of highlighting this tension, he does not tie it in with the theory of evidence that buttresses the EBM movement.

[#]If the goal is to understand treatment decisions as exercises in evidential reasoning, what is not at all clear is what role patient values or their "total well-being" play in engaged decision making. Admittedly, recognition of patient values does shed light of why treatment decisions are often not straightforward mechanical calculations (with patients not finding strong evidence (in the sense of goodness) at all persuasive), but it does at a cost, namely, the patient's sense of their "total well-being" is not conceptualized as a vital part of the patient's evidential reasoning, but is skewed instead as the intrusion of a personal and social variable into a calculation that sometimes results in a bifurcation of values (medical well-being vs total well-being).

(as opposed to the quality of goodness) and thereby the clinician's theoretical and practical reasoning during the clinical encounter.**

Brian Baigrie[1,2]
Mathew Mercuri PhD, Editor-in-Chief[2,3]

[1]*The Institute for the History and Philosophy of Science and Technology, University of Toronto, Toronto, Ontario, Canada*
[2]*The Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada*
[3]*Department of Medicine, McMaster University, Hamilton, Ontario, Canada*

Correspondence
Mathew Mercuri PhD, Assistant Professor, Division of Emergency Medicine, Department of Medicine, McMaster University, Hamilton General Hospital, 237 Barton Street East, Hamilton, Ontario L8L 2X2, Canada.
Email: matmercuri@hotmail.com

ORCID

Mathew Mercuri https://orcid.org/0000-0001-8070-9615

REFERENCES

1. Hacking I. *The Emergence of Probability*. Cambridge: Cambridge University Press; 1975.
2. Cartwright N. Evidence-based policy: What's to be done about relevance? *Philos Stud*. 2009;143:127-136.
3. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence – indirectness. *J Clin Epidemiol*. 2011;64:1303-1310.
4. Straus SE, Glasziou P, Richardson WS, Haynes RB. *Evidence-Based Medicine: how to Practice and Teach it*. 4th ed. Edinburgh: Elsevier; 2011.
5. Djulbegovic B, Guyatt G. Evidence-based medicine and the theory of knowledge. In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd ed. Toronto: McGraw Hill Education; 2015:15, 15-18, 18.
6. Howick J. *The Philosophy of Evidence-Based Medicine*. Chichester, England: Wiley-Blackwell; 2011.
7. Worrall J. What evidence in evidence-based medicine? *Philos Sci*. 2002;69:316-330.
8. Moen J. Evidence-based medicine in context: a pragmatist approach to psychiatric practice. *Philos Psychiatry Psychol*. 2015;22(1):53-62.
9. Buetow S, Kenealy T. Evidence-based medicine: the need for a new definition. *J Eval Clin Pract*. 2000;6(2):85-92.
10. Reiss J. A pragmatist theory of evidence. *Philos Sci*. 2015;82:341-362.
11. Upshur REG. If not evidence, then what? Or does medicine really need a base? *J Eval Clin Pract*. 2002;8(2):113-119.
12. Lipsitch M. Good science is good science. *Boston Rev*. 2020; https://bostonreview.net/science-nature/marc-lipsitch-good-science-good-science.
13. Tufecki Z. What Telling People They Don't Need Masks Backfired. New York Times 2020. https://www.nytimes.com/2020/03/17/opinion/coronavirus-face-masks.html
14. Greenhalgh T. Face coverings for the public: laying straw men to rest. *J Eval Clin Pract*. 2020;26:1070-1077.
15. 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;66:726-735.
16. Neumann I, Santesso N, Akl EA, et al. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol*. 2016;72:45-55.
17. Mercuri M, Upshur REG, Gafni A. Guidelines should not recommend the type of decision-making for the medical encounter. *Patient Educ Couns*. 2020. https://doi.org/10.1016/j.pec.2020.05.016.
18. Mercuri M, Gafni A. The evolution of GRADE (part 1): is there a theoretical and/or empirical basis for the GRADE framework? *J Eval Clin Pract*. 2018;24(5):1203-1210.
19. Mercuri M, Gafni A. The evolution of GRADE (part 2): still searching for a theoretical and/or empirical basis for the GRADE framework. *J Eval Clin Pract*. 2018;24(5):1211-1222.
20. Mercuri M, Gafni A. The evolution of GRADE (part 3): a framework built on science or faith? *J Eval Clin Pract*. 2018;24(5):1223-1231.
21. Veatch RM. Doctor does not know best: why in the new century physicians must stop trying to benefit patients. *J Med Philos*. 2000;25(6):701-721.
22. Angell M. The legacy of Karen Ann Quinlan. *Trends Health Care Law Ethics*. 1993;8:17-19.

**Clinicians, Veatch[21(p705)] reminds the reader, are well-versed in medical conceptions of well-being. The preservation of life is built into the Hippocratic Oath and, in the wake of widely publicized legal cases; for example, Karen Quinlan (see[22]), clinicians have been under mounting pressure to adopt a more pluralistic position on conceptions of medical well-being. It is also worth noting that as medicine has become increasingly specialized, clinical specialties have emerged that are tailored to deal with just one of the many formulations of medical well-being (eg, relief of pain and suffering, saving life).