REVIEW

THE AMERICAN JOURNAL of MEDICINE®

From Evidence Based Medicine to Medicine Based Evidence

Ralph I. Horwitz, MD, MACP,[a] Allison Hayes-Conroy, PhD,[b] Roberto Caricchio, MD,[c] Burton H. Singer, PhD[d]

[a]Temple Transformative Medicine Institute; [b]Geography and Urban Studies; [c]Temple Lupus Clinic, Temple University, Philadelphia, Pa; [d]Emerging Pathogens Institute, University of Florida, Gainesville, Fla.

ABSTRACT

Evidence based medicine, using randomized controlled trials and meta-analyses as the major tools and sources of evidence about average results for heterogeneous groups of patients, developed as a reaction against poorly designed observational treatment research and physician reliance on personal experience with other patients as a guide to decision-making about a patient at hand. However, these tools do not answer the clinician's question: "Will a given therapeutic regimen help my patient at a given point in her/his clinical course?" We introduce fine-grained profiling of the patient at hand, accompanied by comparative evidence of responses from approximate matches to this patient on whom a contemplated treatment has/has not been administered. This represents medicine based evidence that is tuned to decision-making for the particular patient.

 • The American Journal of Medicine (2017) 130, 1246–1250

KEYWORDS: Approximate match; Decision-making; Patient profiling; Personalized evidence

SEE RELATED ARTICLE, p. 1290.

INTRODUCTION

Evidence based medicine developed as a reaction against poorly designed observational treatment research and physician reliance on personal experience with other patients. Randomized controlled trials (RCTs) and meta-analyses are the major tools of evidence based medicine and the source of the evidence that describes average results for groups of patients. Emphasizing highly selected RCT populations, placebo controls, and "hard" endpoints (death and major morbidity), evidence based medicine became established as the scientific basis for Population Medicine.[1]

Evidence based medicine has created enormous benefits for population health. Average results from traditionally designed RCTs that emphasize internal validity (to minimize bias in design) over generalizability (to apply the results to the patients who will use the therapies) is exactly what is needed for pharmaceutical companies who are developing drugs and regulators who license them for use in the population.[2] By separating useful from useless therapies, evidence based medicine has been instrumental in providing the evidential basis for effective population-level control of risk factors for myocardial infarction and stroke, has played a critical role in the transformation of HIV from a fatal infection to a chronic disease, was instrumental in testing drugs that can now cure hepatitis C virus in many patients, and has been the basis for rigorous verification of substantial improvements in the outcomes of some cancers.

These achievements notwithstanding, limitations of evidence based medicine are now increasingly evident. Narrow criteria for inclusion in RCTs frequently exclude the very patients who will use the medicine after it receives regulatory approval (eg, it is estimated that studies of medications for asthma have excluded 95% of the target population).[3] The use of placebo controls may exaggerate treatment benefits especially when, as is often the case, new drugs are not tested against effective comparative therapies. The routine reliance on "hard" end points in RCTs disregards the many outcomes of treatment such as physical, social, and psychological well-being that are deeply important to patients.[4,5]

The limitations of the RCT, as currently utilized in clinical medicine, are acknowledged by many of its advocates. Efforts to mitigate the negative effects are increasingly

Funding: None.

Conflict of Interest: None for all authors.

Authorship: All authors had access to the data and a role in writing the manuscript.

Requests for reprints should be addressed to Ralph I. Horwitz, MD, MACP, Temple Transformative Medicine Institute, 3500 North Broad Street, MERB, Room 1111E, Philadelphia, PA 19141.

E-mail address: Ralph.horwitz@temple.edu

0002-9343/$ - see front matter
https://doi.org/10.1016/j.amjmed.2017.06.012

implemented. For example, the US Food and Drug Administration has encouraged trialists to broaden the populations studied in RCTs.[6,7] Pragmatic RCTs have emerged as an alternative to traditional RCTs to emulate more closely the actual practice of medicine and foster more comparative effectiveness studies. An emphasis on more comprehensive studies of patients has provided impetus to develop indexes and scales that measure patient-centered outcomes such as well-being and physical and social functioning.[8] Despite this progress, none of these efforts are able to remedy the most fundamental limitation of evidence based medicine. It provides only a coarse-level population-based model of clinical care. It is ill suited to providing the personalized evidence that the clinician needs to guide the care of an individual patient. Clinical medicine has always been focused on the individual patient, and the weakness of the RCT (and later, evidence based medicine) to guide physician decision-making was recognized early by those who developed the method and knew it best. In his Heberden Oration in 1965, Austin Bradford Hill wrote, "This leads to a related criticism of the present controlled trial that it does not tell the doctor what he wants to know. It may be so constituted as to show without any doubt that treatment A is on the average better than treatment B. On the other hand, that result does not answer the practicing doctor's question*: what is the most likely outcome when this particular drug is given to a particular patient?"* (italics added).[9]

CLINICAL SIGNIFICANCE

- Medicine based evidence starts with a longitudinal profile of the biological, clinical, psychological, and social environmental history of a single index patient.
- Profiles that approximately match the index patient provide the comparative empirical base for management of the index case.
- The evidential core of medicine based evidence, the approximate matches to an index case, is focused on the needs of clinical practice and sharply at odds with evidence based medicine.

EVIDENCE BASED MEDICINE AND THE ROLE OF SUBGROUPS

Clinicians have long been troubled by this shortcoming of RCTs. One strategy for addressing Hill's[9] challenge has been for clinicians to request an ever-increasing number of subgroup analyses to supplement the overall results of a trial. Resistance by trialists to the clinician's frequent request for sub-groups led to a tongue-in-cheek analysis by the investigators in the ISIS-2 trial.[10] In that study, the analysts looked at 12 subgroups formed by the 12 signs of the Zodiac. When 2 of the subgroups (Libra or Gemini) had differences that were discordant from the trial's overall average results, the authors concluded that subgroups are prone to misleading results by chance alone. The absence of more attention to subgroups created by relevant biological, clinical, and social features represents a serious deficiency in current clinical science. Importantly, the subgroups need to be formed not just according to static baseline features but also by dynamic changes in clinical and psychosocial features that change over the course of the trial.

From the perspective of the clinician focused on management of an individual patient, specification of a sub-group of interest within a large RCT is equivalent to asking for approximate matches to the patient at hand. A fundamental point is that the subgroup cannot be defined until there is an index case patient to whom the subgroup will be compared. The response to treatment(s) of the approximate matches provides guidance for what may be anticipated if the same treatment is applied to the individual patient.

THE INCOMPATIBLE TWO-HORSE TEAM

Most subgroup analyses in RCTs are based on variables measured at baseline when patients are randomized to one or the other treatment arm. In constructing subgroups, the investigators typically rely on a single variable, such as sex, age, or levels of disease severity. Occasionally, such an analysis uncovers a previously undiscovered treatment effect, as, for example, when a subgroup of patients with left main coronary artery disease had benefits from coronary artery bypass grafting[11] that were not present in the negative results of the overall study population. Recently, it has been pointed out that men and women may respond differently to drug therapies. According to one group of scientists, "Aspirin for the primary prevention of cardiovascular disease is one such example. A meta-analysis of 6 trials found that aspirin in men reduced the risk of myocardial infarction by 32% but had no effect on ischemic stroke. In contrast, aspirin in women had no effect on myocardial infarction but reduced the risk of stroke by 24%".[12]

Subgroup analyses that rely on single variables measured at baseline fail to represent the complexity of clinical medicine where patients have multiple features that need to be measured both at baseline and in follow-up examinations that replicate more closely how patients are monitored and managed by physicians in practice. Single variable subgroup analyses have another liability: the multiplicity problem that arises when studies report the results of multiple statistical tests of subgroups raising the probability that at least some of the results will be found to be statistically significant even if there is no underlying effect.

One of the great American statisticians, John Tukey, discussed the multiplicity problem in a famous 1979 paper, but also used the opportunity to refer to what he perceived to be an even more intractable problem:[13] "It is a difficult task to drive the nearly incompatible two horse team: On the one hand, knowledge of a most carefully evaluated kind where in

particular questions of multiplicity are faced up to, and on the other, informed professional opinion, where impressions gained from statistically inadequate numbers of cases often, and so far as we can see, should control the treatment of individual patients. The same physician or surgeon must be concerned with both what is his knowledge and what is his informed professional opinion as part of treating a single patient. I wish I understood better how to help in this essentially ambivalent task."

CLINICIANS' DILEMMA AND THE RISE OF PERSONALIZED MEDICINE

The concerns of Bradford Hill and Tukey have framed the dilemma faced by every physician caring for an individual patient.[9,13] The ideal information base for a clinician managing the ongoing clinical course of a patient is not the traditional RCT that forms treatment categories at baseline by randomized assignment and then waits until the end of the trial to count outcomes, ignoring everything that has occurred in between in the care of the patient. Consider the Women's Health Trial that randomized 16,000 postmenopausal women to either hormone replacement therapy (HRT) or placebo and planned an 8-year follow-up to test whether HRT would improve certain clinical outcomes. The study was stopped early when an increased risk was observed for coronary heart disease among women assigned to receive HRT.

The interpretation of these results was complicated when it was noted that treatment was discontinued and blinding was broken for nearly half of the HRT users but only a small percentage of the placebo users.[14] The changes in treatment occurred mostly to manage vaginal bleeding that developed as the trial progressed among women assigned to estrogen therapy. But the loss of blinding created unanticipated problems in detection bias and in the adjudication of end points (because patients and physicians were unblinded by a treatment side effect), and the changes in treatment assignment meant that the customary intent-to-treat analysis no longer could answer the question that had been originally proposed. Instead of answering a question on the efficacy of estrogens on cardiovascular disease, the study was now a test of initiating a treatment that many patients could not maintain and had changed during follow-up.

What are needed instead to guide management of individual patients are case records that describe the actual clinical course and treatment management decisions by physicians. In this regard, what was discovered in the Women's Health Trial when vaginal bleeding was further investigated? How did discontinuation of estrogen affect cardiovascular risk management? Were medications added to intensify control of blood pressure, hyperlipidemia, glucose regulation, or other cardiovascular risk factors? Were patients advised to exercise more or adjust their diets? When did these first-order treatment modifications occur and what were the intermediate outcomes?

To answer these questions for a particular patient and provide comparative evidence from a wider patient population requires an archive of clinical histories closely matched to the patient in question. Some of the individuals in the comparison population will have received the contemplated treatment including follow-on treatment modifications and others will not. Information on the clinical course of the patient at hand is monitored, recorded, and analyzed comparatively with the same, or closely approximated, information from patients in the comparison archive. The fundamental point is that the relevant comparison histories approximate the clinical practice details of the patient at hand. Medicine based evidence like that just described is critically needed if we are to realize the personalized care of patients that is the promise of scientific advances in biomedicine.

OVERCOMING THE BARRIERS TO PERSONALIZED MEDICINE

The rise of evidence based medicine was fostered by the concern that treatment evaluations using observational study designs were biased if randomization was not used to assign treatment. This argument was true in historical controlled trials when patients receiving a new treatment were compared with historical patients from a different secular era.[15] However, as always happens in science, improvements were made to the design and analysis of observational study designs so that the average results of RCTs and the average results of observational studies of the same treatment now produce similar results.[16,17]

The rejection of physician experience, what Tukey refers to as "informed professional opinion," was also warranted when the experience was limited to a single doctor (experience with a little "e"). Now, however, advances in computing and informatics makes it possible to access and analyze the collected experience of tens of thousands of physicians caring for hundreds of thousands of patients (experience with a big "E"), far more than could ever be enrolled in a single clinical trial, thus reducing the worry that treatment decisions are based on selected physician experience. It is true that the analysis of aggregate physician experience will identify wide variation in clinical practice. This heterogeneity of practice patterns is an advantage of this approach because it enables consideration of patients' clinical courses under diverse treatments and for patients with diverse histories. The central point is that now our considerations are focused entirely on issues of clinical practice, and this is where analysis to guide decision-making regarding an individual patient must be centered.

Medicine based evidence will require establishment of an archive of profiles using information on clinical experience from all types of studies: observational cohorts and case control studies, traditional RCTs and pragmatic RCTs, real-world studies and disease registries will all contribute to the needed database.[18] Equally important, medicine based evidence will incorporate both biological and biographical (life experience) information to suitably profile patients. The inclusion of biographical information is consistent with wider trends in biomedicine and the life sciences, which have acknowledged that a broad array of both pivotal

and recurrent life events may greatly shape health and disease.

A major shift in the specification of evidence to guide clinical practice has been fostered by recent scientific and technological advances in genetic and molecular science that are unraveling disease pathogenesis at the level of the individual. One exciting example of this development is the recent report of the "narciss-ome," one person's integrative personalomics profile created by merging the genomic sequence with RNA, protein, metabolic, and auto-antibody profiles measured 20 times over a 14-month period.[19] Not only were new disease susceptibilities identified, but the patient's blood glucose levels escalated following a viral infection, indicating a diabetic state that was only diagnosed later. And already, clinical medicine is fast developing new ways to profile patients. Current devices are capable of "digitizing" a person with wearable sensors that quantify physiological metrics such as vital signs, provide high-definition images of a persons' anatomy, and characterize the microbiome. An analogous technology has been developed to utilize biographical information in narrative form and succinctly incorporate it in a more nuanced patient profile than is currently customary. In addition to the customary measurements of genomics and metabolomics, we will also need to include the exposome that is defined as the measure of all the exposures of an individual in a lifetime, beginning before birth and including environmental and occupational sources.

BACK TO THE FUTURE: MEDICINE BASED EVIDENCE PAST, PRESENT, AND FUTURE

We would like to believe that the strategy of medicine based evidence that we propose is new, but surprisingly there are footprints of this approach in previous research. Feinstein et al imagined a "library of clinical experience" to obtain personalized prognosis.[20] Indeed, in 1972 Feinstein et al[20] proposed an early illustration of the process we have outlined and did so with far more limited computer capabilities than are currently available. The basic idea of starting with a profile of an individual patient and identifying approximate matches to the patient in a library of clinical experience and using these cases as a guide to decision-making for the patient at hand is in this visionary paper. In the short run, there is a pressing need to develop algorithms for approximate matching and to do so on patients with a diversity of diseases. The matching would be based on much more comprehensive information than in the original example of Feinstein et al,[20] but their fundamental philosophy about the character of the evidence needed to guide decision-making for individual patients is salient today.

One of the barriers to developing medicine based evidence is the ideological adherence to evidence based medicine as the only "scientific" approach to clinical decision-making. For example, evidence based medicine requires investigators to design studies that prefer internal validity to external validity, when the goal of clinical research should be results that are accurate for all those who might receive the intervention, those included in the trial and those outside the trials.[21]

WHAT MEDICINE BASED EVIDENCE IS AND ISN'T

Each time the physician considers a change in the management of an individual patient, she is engaged in an act of clinical prediction. The information used for this prediction consists of the full history of the index case up until a time when an intervention is contemplated and augmented by the experience with and without the contemplated intervention(s) on approximate matches to the index case. In fact, the physician will likely contemplate and carry out interventions at multiple times during the clinical course of the index case. Thus, guidance from sets of approximate matches, who are not necessarily the same set of cases at each intervention time, is precisely what is needed to enhance the information base for management of the index case over time.

Having described the clinical prediction problem, it is also important to say what is not a focus of medicine based evidence. While assembling the evidence for clinical prediction, we are not testing the efficacy of a contemplated intervention. We are using what may be the results of such tests on the approximate matches to the index case. But, medicine based evidence is not an N = 1 trial. We are not interested in knowing what the contemplated intervention is likely to do across a range of patients whose clinical courses prior to the intervention time are quite different from that of the patient at hand. We are also not interested in knowing how the contemplated intervention acts holding levels on multiple other variables constant. We want insight about the intervention's likely performance on patients whose complex profile evolving over time is like the multidimensional longitudinal profile of the patient whose clinical care we are managing. What happens to the patient before us is of paramount importance. What might happen with other kinds of patients is of no concern.

The US Congress recently passed the 21st Century Cures Act that was strongly supported by the Director of the Food and Drug Administration Center for Drug Evaluation and Research, who commented: "Patient-focused drug development has the potential to become the most transformational piece of this bill."[22] What she was referring to was the opportunity to collect structured data from a broad range of patients and to use those data to help guide physician and patient decision-making. Neither the "moonshot" for cancer[23] nor the evidential base for personalized (or precision) medicine will be successful if we use the old methods of evidence based medicine. The time is now for innovation in translational science by developing and implementing medicine based evidence to provide the right patient with the right treatment at all the right times.

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312:71-72.

2. Horwitz RI, Singer BH, Makuch RW, Viscoli CM. Can treatment that is helpful on average be harmful to some patients? A study of the conflicting information needs of clinical inquiry and drug regulation. *J Clin Epidemiol*. 1996;49:395-400.
3. Herland K, Akselsen JP, Skjønsberg OH, Bjermer L. How representative are clinical study patients with asthma or COPD for a larger "real life" population of patients with obstructive lung disease? *Respir Med*. 2005;99:11-19.
4. Rawlins MD. Harveian Oration. De testimonio: on the evidence for decisions about the use of therapeutic interventions. *Lancet*. 2008;372:2152-2161.
5. Fava GA, Guidi J, Rafanelli C, Sonino N. The clinical inadequacy of evidence-based medicine and the need for a conceptual framework based on clinical judgment. *Psychother Psychosom*. 2015;84:1-3.
6. Temple R. A Good Review Practice MAPP [Clinical Review of Investigational New Drug Applications] describes our obligation to consider breadth of patient populations in trials at EOP2 and other meetings during drug development and to discourage unneces- sary exclusions of patients. Memorandum to OND/OTS Review Staff. Dec 2, 2013.
7. Roland M, Torgerson DJ. Understanding controlled trials: what are pragmatic trials? *BMJ*. 1998;316:285.
8. Cella D, Yount S, Rothrock N, et al. The patient-reported outcomes measurement information system (PROMIS). *Med Care*. 2007;45:S3-S11.
9. Hill AB. Reflections on controlled trial. *Ann Rheum Dis*. 1996;25:107.
10. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. *Lancet*. 1988;332:348-360.
11. Chaitman BR, Fisher LD, Bourassa MG, et al. Effect of coronary bypass surgery on survival patterns in subsets of patients with left main coronary artery disease. Report of the Collaborative Study in Coronary Artery Surgery (CASS). *Am J Cardiol*. 1981;48:765-777.
12. Berger JS, Roncaglioni MC, Avanzini F, Pangrazzi I, Tognoni G, Brown DL. Aspirin for the primary prevention of cardiovascular events in women and men: a sex-specific meta-analysis of randomized controlled trials. *JAMA*. 2006;295:306-313.
13. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science*. 1977;198:679-684.
14. Garbe E, Suissa S. Issues to debate on the Women's Health Initiative (WHI) study: hormone replacement therapy and acute coronary outcomes: methodological issues between randomized and observational studies. *Hum Reprod*. 2004;19:8-13.
15. Diehl LF, Perry DJ. A comparison of randomized concurrent control groups with matched historical control groups: are historical controls valid? *J Clin Oncol*. 1986;4:1114-1120.
16. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342:1887-1889.
17. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342:1878-1886.
18. Horwitz RI, Singer BH. Why evidence based medicine failed in patient care and medicine based evidence will succeed. *J Clin Epidemiol*. 2017;84:14-17.
19. Dennis C. The rise of the "narciss-ome". *Nature*. 2012;doi:10.1038/nature.2012.10240.
20. Feinstein AR, Rubinstein JF, Ramshaw WA. Estimating prognosis with the aid of a conversational-mode computer program. *Ann Intern Med*. 1972;76:911-921.
21. Concato J, Horwitz RI. Beyond randomized versus observational studies. *Lancet*. 2004;363:1660-1661.
22. Woodcock J. 21st century cures act: pork or promise? *Forbes*. 2016. Available at: https://www.forbes.com/sites/judystone/2016/09/22/21st-century-cures-act-pork-or-promise/.
23. Horwitz RI. Equity in cancer care and the outcomes of treatment: a different type of cancer moonshoot. *JAMA*. 2016;315:1231-1232.