ESSAY REVIEW

Research and medicine: human conjectures at every turn

Eyal Shahar MD MPH

Professor, Division of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona, USA

Keywords
Absolute risk, authoritarianism, causality, epidemiology, evidence-based medicine, person-centered medicine, probability, relative risk, statistics-based medicine

Correspondence Address
Eyal Shahar, Division of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, The University of Arizona, 1295 N. Martin Avenue, Tucson, AZ, USA. E-mail: shahar@email.arizona.edu

Accepted for publication: 17 May 2011

Do we know anything?

In his book, provocatively titled 'stats.con' (not 'com'), Penston examines the sources of contemporary medical knowledge and attacks them all: randomized trials, observational studies, schools of statistics and more [1]. His text covers a wide range of topics, some of which are as philosophical as determinism and indeterminism and others are as practical as reports of fraudulent studies. He has done a fair amount of homework and I found myself agreeing with his arguments many times - although not always with his conclusions. Whether I agreed with him or not, it was refreshing to encounter a courageous author who bluntly states what he thinks. Penston writes passionately and clearly, which makes it easy to identify both mistakes and points of disagreement. We need more physicians who genuinely care about epistemology and methodology and don't just look for a magical *P*-value.

Penston sets the stage for his book in the first paragraph of the Preface. He writes: "Nowadays, much of what passes for knowledge is derived from statistics-based research." Then, a few sentences later he asks: "But how confident should we be about this type of evidence?" Presumably, the confidence has to do with a medical decision for Mr. Smith across the desk or for Ms. Jones in room 7A across the nurses' station. If so, I would like to offer my answer: We should have no confidence in any type of evidence - whether it comes from research with statistics, from research without statistics, or from years of clinical experience. I repeat: We should have no confidence at all, for it is always possible that a single decision for a single patient will be plainly wrong [2]. Those who are interested in being confident need to study the causal process by which a psychological state of confidence is reached. Of course, the question remains: how confident should they be about the evidence?

To show my students the conjectural nature of medical knowledge, I sometimes ask them the following question: "How can we know that the main result of a mega-trial was not fabricated?" After some thinking, several students usually suggest appointing an external committee that would scrutinize the study and issue a verdict. I continue: "And how can we know that the verdict was honest?" My audience usually remains silent and I suggest a second committee that would scrutinize the first, and then a third committee that would scrutinize the second and so on. At this point, the students are split into two groups: those who think that I am just an annoying skeptic, and those who understand the difference between human belief and the conjectural nature of knowledge [3,4]. So let's make it clear at the outset. Medical research does not generate valid knowledge, true knowledge, confident knowledge, justified knowledge, reasonable knowledge, or anything of the kind. Medical research, like all scientific research, generates *conjectural* knowledge, which always remains open to revisions [3]. Then, all kinds of social forces periodically turn conjectural knowledge into boundaries of acceptable conduct, definitions of malpractice, and so on.

On the purpose of medical research

If I were the editor of the book I would have placed two late chapters at the very beginning. The chapters titled 'Hume *et al*.' and 'Imperfect Causal Inference' should be the starting point. They deal with causation, which is the essence of all of science, including medical research. It is difficult to appreciate a critique of methodology before we know the purpose of the inquiry: what do we want to

know? Penston correctly states early on that the knowledge we seek is causal: "…the purpose of research is to identify causal relationships". Unfortunately however, he chose the language of null hypothesis testing, which he later attacks with full force. To identify a causal relation is to issue a verdict about whether *A* causes *B*, or to state a belief on the matter. In my view, this is not the purpose of medical research at all. And this argument alone renders null hypothesis testing irrelevant in any of its versions [5] and amalgamations [6].

In my view, the purpose of medical research is to estimate the magnitude of the effect of a causal contrast, for example the probability ratio of a binary outcome (*B*) for the contrast between taking some drug and not taking it (*A*). In notation: $\Pr(B=1|A=1)/\Pr(B=1|A=0)$ and $\Pr(B=0|A=1)/\Pr(B=0|A=0)$. That's what we want to know about causal reality. If the ratios are close to 1, the effect is weak, and possibly uninteresting, even if it is not precisely null. If the ratios are far from 1, the effect is strong and usually interesting. The underlying axiom states that Nature not only decided that *A* would affect *B*, but also decided how strong of an effect *A* would have on *B*. Furthermore, Nature has decided whether that effect would vary according to the values of some other variable(s)[7,8].

The probabilities in question are manifestation of indeterminism [9], which Penston and others call probabilistic causation. Penston criticizes the idea on several counts – unjustly in my opinion. I can offer here only a brief rebuttal. First, in probabilistic causation, unlike deterministic causation, there are no causes of any event. For example, it is wrong to say that smoking has caused someone's death or that taking a drug has saved someone's life. Rather, there are causal variables and effect variables: smoking status (or drug taking status) is a cause of vital status. Second, a cause does not 'increase the probability of its effect'. Rather, the probability of the effect variable varies according the value of the causal variable: 'up' or 'down' is just an arbitrary choice of which value serves as the reference. Third, reversed probability, which Penston lists as another difficulty of probabilistic causation, is not manifestation of causality. If *A* affects *B*, then $\Pr(A=1|B=1)/\Pr(A=1|B=0)$ is not a causal probability ratio; it is just math. Moreover, the math of probabilities should be viewed as a fallible, human-made method to study indeterministic causal propensities (or tendencies), not as reality itself. Nature did not need to compute any numerical probabilities in order to create an indeterministic causal structure. We - human beings - assume that mathematical ideas can help us to discover the secrets of Nature.

Penston thinks that many people believe in probabilistic causation, or as I prefer to call it—indeterminism. I wish he were right. My observations suggest that most scientists (and laypersons, too) repeatedly switch between deterministic claims and indeterministic claims. On the one hand, they might say that smoking has caused that many cases of lung cancer, which is a classical expression of determinism. On the other hand, they might also say that your chances of getting lung cancer are higher if you smoke than if you don't smoke, which is a classical expression of indeterminism. Moreover, not very many scientists understand the strong connection between research methodology and the axiomatic choice between determinism and indeterminism [2,8].

All kinds of statistics

Penston criticizes 'statistics-based research in medicine'. But what exactly is statistics-based medical research? Here, I think he might have taken a wrong path, combining too many methods under a single title: Does the study design belong to statistics or to research methodology? Are all statistical methods at fault? Did he really wish to argue against every aspect of quantitative science? I hope not.

Rather than counting points of disagreement, I will simply share my thoughts on several topics that Penston has examined. You might wish to consider them while reading his book.

First, a causal analysis of a study should generate two numbers: a measure of association and a standard error. The authors should consider the former as a *conjectural* measure of effect. If they don't, they have nothing interesting to share. The standard error is some function of the sample size, and should be considered as one of many features of the study. I fully agree with Penston that attempts to extract anything else from the estimate and the standard error do not generate any new knowledge. That includes every version of null hypothesis testing (Fisherian, Neymanian, Bayesian), confidence intervals of the frequentist, and credible intervals of the Bayesian. Yes, we have all been fooled by these statistics for many years, and it's time to leave them behind, including the Bayesian approach for which Penston shows some sympathy. On the other hand, I don't see anything wrong in regression models and likelihood functions, which are statistical means to estimate the effect of interest. Again, I don't have confidence in any number on a printout, but the maximum likelihood estimate of the rate ratio from a Poisson regression model does generate conjectural knowledge. Not certain knowledge, not reliable knowledge, but knowledge nonetheless. Perhaps even Penston would agree.

Second, once we ignore the uninteresting question 'does *A* affect *B*?', it should be obvious that all research results are wrong. Only miraculously would the estimate from a study be identical to the causal parameter [8]. Therefore, the scientific debate should not center on whether the estimate is right - for it is not - but on the method that generated the estimate. As difficult as it may be to accept, we have to comply with the following rule of science: You may argue feverishly about the method behind an estimate, but you should remain silent about the estimate itself. Ironically perhaps, card-carrying statisticians know this rule better than others and many of them adhere to it, too. For instance, thoughtful statisticians might praise an *estimator* (i.e., the method that generated an estimate) for being unbiased, but they would not attach

this adjective to the *estimate* itself. They know that the term 'unbiased estimate' is a misnomer. They know that the scientific method cannot tell us anything about any estimate.

Finally, the *estimator* from any study, including a randomized trial, is prone to six types of bias: confounding, colliding bias, information bias, thought bias, effect-modification bias, and causal pathway bias [8, 10-14]. Many examples in Penston's book nicely fit this taxonomy, which is founded on the principles of causal diagrams. Nonetheless, there is a big difference between understanding the taxonomy and using it to criticize any specific study. Yes, the estimator from my cohort study might contain confounding bias, but until you propose an explicit confounding theory, you are not engaged in a scientific exchange [15]. You are just reiterating the obvious: all scientific knowledge is fallible knowledge. Philosophers have known it for many years, but medical researchers have yet to catch up. Moreover, even if my estimator were bias-free, my estimate is still just one number from an entire distribution (assuming indeterminism) and neither you nor I can say anything about the distance between my estimate and the causal parameter. Nor can statistics tell us whether any observed difference (i.e., an estimate) is due to chance. It is all wishful thinking, erroneous beliefs, or misunderstanding of statistics.

Authoritarianism

Physicians want to know "whether the findings are applicable to that particular individual", or as I would rephrase: what is the effect of variable *A* on variable *B* in that patient? I am sorry to say again that no one holds the answer or will ever hold it [16,17]. Causation, whether deterministic or indeterministic, is not directly observable and cannot be deduced from a set of certain premises. We are doomed to remain ignorant about causal relations, even when we strongly believe otherwise.

Where certain knowledge cannot exist, various forms of authoritarianism quickly fill the void. Not that long ago, paternalistic medicine was the norm: the doctor was the authority and the patient was supposed to follow orders quietly and respectfully. There are still many doctors who hold that view, but social forces have empowered patients with authority, too. It is called 'patient autonomy'.

I think that Penston would have endorsed a model of shared authority between a patient and his or her doctor. Many patients don't want full authority over their care, because it might defy the very purpose of medical care. I want to feel better, even if the means is the placebo effect, which my doctor can control, unbeknown to me. Moreover, authority comes with mental burden and many patients would gladly transfer some of that burden to their doctor. I guess we may still call it 'patient autonomy': the autonomy to give up authority.

Sadly, things have drastically changed in medicine (and elsewhere). As Penston writes, new authorities and new forms of authoritarianism have emerged:

"In recent times, there's a strong impression of authoritarianism creeping into our lives…Trust in individuals is no longer enough. There must be constant surveillance to ensure that they're acting in line with the current orthodoxy. The freedom of professionals to take decisions based on their learning and experience is shunned, and strict rules are put in place which, if not adhered to, lead to censure."

Indeed, others have joined the special encounter between a sick person and his or her doctor. They come from numerous corners of our society: government officials, members of consensus-generating committees, authors of practice guidelines, followers of the evidence-based medicine crusade, salary-paid ethicists, and others. They were not invited by the patient or his doctor, but they showed up anyway. Who let them in? How did they justify their intrusion into a private medical encounter?

The answer is simple. The intruders claim to have authority over knowledge and sometimes over moral principles as well [18,19]. They know how to turn an estimated effect and a standard error into something better. I just wish they had told me their secret: how do they know that which cannot be known?

According to Penston, statistics-based research is the culprit, having been "a crucial weapon in the new authoritarianism". But in my view, he is blaming the weapon for the wrongdoing of the perpetrator. Medical research is a weapon in our never-ending fight against ignorance. We should relentlessly try to fix it, as best as we can, rather than discard it all. I would gladly join the fight against *P*-values and confidence intervals [20,21], but would stand up to defend case-control studies and logistic regression.

We should fight the intruders and their false claims of superior knowledge on three fronts, neither of which requires us to abandon all research methods. First, attack bad statistics as Penston has done: show that key statistical practices, on which the intruders rely, are irrelevant, or misleading, or plainly wrong. Second, repeatedly remind the intruders that there are no authorities over knowledge. Any scientific assertion about causal reality is a conjecture, no matter how many people line up behind it. Third, emphasize that medical practice, just like public health practice [22], relies on science, but should not be equated with science [23]. The medical encounter contains a lot more than research-based algorithms.

Person-centered medicine

The name of this journal says it all. If you had proposed a journal by this name 50 years ago, no one would have understood what you had in mind. The patient was located at the center of medical practice and the doctor did whatever he could to cure, help, comfort and be present when nothing more could be done. These days the patient is circling

other parties: the resident who decides if the patient is 'pleasant' and 'compliant'; the specialist who counts 'minutes of consultation'; the insurer who approves (or not) a request for a diagnostic procedure; and the hospital regulator who collects your signature ten times before you are admitted. How did we let it happen? How did medicine become a business-oriented, depersonalized, and algorithmic practice? Why does 'care giver' sound like the opposite of someone who cares? Some day, sociologists and historians might provide the answers.

The battle to reclaim humanistic medicine might be lost, but even lost battles should be fought. We should not lose hope as long as there are people who are willing to fight for a good cause. Parts of Penston's book [1] and this new journal [24] should serve us well on the battle field.

References

[1]. Penston, J. (2010). stats.con: How we've been fooled by statistics-based research in medicine. The London Press, UK.

[2]. Shahar, E. (2008). Does anyone know the road from a randomized trial to personalized medicine? *Journal of Evaluation in Clinical Practice* 14, 726-731.

[3]. Popper K.R. (1959). The Logic of Scientific Discovery. Basic Books, New York.

[4]. Miller, D. (1994). Conjectural knowledge. In Critical Rationalism: a Restatement and Defence, pp. 1–13. Open Court, Chicago.

[5]. Oakes, M. (1990). Statistical Inference. Epidemiology Resources Inc, Chestnut Hill.

[6]. Hubbard, R. and Bayarri, M. J. (2003) Confusion over measures of evidence (p's) versus errors (α's) in classical statistical testing. *The American Statistician* 57, 171–178.

[7]. Shahar, E. and Shahar, D.J. (2010). On the definition of effect modification. *Epidemiology* 21, 587.

[8]. Shahar, E. (2007). Estimating causal parameters without target populations. *Journal of Evaluation in Clinical Practice* 13, 814-816.

[9]. Popper K.R. & Bartley W.W. (1982) The Open Universe: an Argument for Indeterminism. Rowman & Littlefield, Totowa, NJ.

[10]. Pearl, J. (2009) Causality: Models, Reasoning, and Inference. 2nd ed. Cambridge, UK: Cambridge University Press.

[11]. Hernan, M.A., Hernandez-Diaz, S. and Robins, J.M. (2004) A structural approach to selection bias. *Epidemiology* 15, 615–625.

[12]. Shahar, E. and Shahar, D.J. (2009). On the causal structure of information bias and confounding bias in randomized trials. *Journal of Evaluation in Clinical Practice* 15, 1214-1216.

[13]. Shahar, E. and Shahar, D.J. (2010). Causal diagrams, information bias, and thought bias. *Pragmatic and Observational Research* 1, 33–47.

[14]. Shahar, E. and Shahar, D.J. (2010). More on selection bias. *Epidemiology* 21,429-430.

[15]. Shahar E. A method to detect an unknown confounder: something from nothing? *Journal of Evaluation in Clinical Practice*, Epub March 13, 2011

[16]. Shahar, E. (1997). Translating clinical trials into practice. *Lancet* 349, 654-655.

[17]. Shahar, E. (1998). Evidence-based medicine: a new paradigm or the Emperor's new clothes? *Journal of Evaluation in Clinical Practice* 4, 277-282.

[18]. Miles A., Grey J., Polychronis A. and Melchiorri C. (2002) Critical advances in the evaluation and development of clinical care. *Journal of Evaluation in Clinical Practice* 8, 87–102.

[19]. Shahar, E. (2003). On morality and logic in medical practice: commentary on 'A critical appraisal of evidence-based medicine: some ethical considerations'. *Journal of Evaluation in Clinical Practice* 9, 133-135.

[20]. Shahar, E. (2007). Commentary: interpreting the interpretation. *Journal of Evaluation in Clinical Practice* 13, 693-694.

[21]. Poole, C. (1987) Confidence intervals exclude nothing. *American Journal of Public Health* 77, 492–493.

[22]. Shahar, E. (1996): Epidemiology differs from public health practice. *Epidemiology* 7, 558-559.

[23]. Miles, A. (2009). On a Medicine of the Whole Person: away from scientistic reductionism and towards the embrace of the complex in clinical practice. *Journal of Evaluation in Clinical Practice* 15**,** 941–949.

[24]. Miles, A. and Mezzich, J.E. (2011). Editorial introduction: Advancing the global communication of scholarship and research for personalized healthcare: The International Journal of Person Centered Medicine. *International Journal of Person Centered Medicine* 1(1), 1-5.