Medicine—turning on its head!

Om Prakash Yadava¹

Received: 16 April 2019 / Accepted: 3 May 2019 / Published online: 17 May 2019
© Indian Association of Cardiovascular-Thoracic Surgeons 2019

Three totally unrelated and stand alone publications [1–3], behove us to try and answer the conundrum—how scientific is science? Let me break it up.

How guiding are the guidelines?

Fanaroff et al. report that more than 90% of recommendations in the guidelines from the American College of Cardiology (ACC), American Heart Association (AHA) and European Society of Cardiology (ESC) were based on weak evidence. In 26 current ACC/AHA guidelines, only 8.5% of the recommendations were classified as level of evidence (LOE) ‘A’. It was slightly better with 25 of the current ESC guidelines, where 14.2% recommendations were having LOE ‘A’ [1]. What is more glaring and disconcerting is the fact that another study, a decade ago by the same group, had found nearly similarly findings and there had been no progress at all. In fact, ‘if any thing, the situation has gotten slight worse’ [1]. Should guidelines thus be eschewed? My take—no; yet, the guidelines are very much needed. In an accompanying editorial, Bonow and Braunwald comment, ‘although guidelines are imperfect and a work in progress, they remain the corner stone for informing clinical decisions’ [4]. We could not agree more as long as one does not take these guidelines as mandatory, but only clinically directive. Additionally, these should adapt to the ground realities, factoring in logistics, cultural diversity, local practices and beliefs and must answer the aspirations of the society [5].

Is ‘*p*’ value the holy grail?

Even if the guidelines were to be based on randomised controlled trials, thereby improving the LOE, the very basis of the validity of these trials has been questioned, as the authenticity of the use of ‘*p*’ value, on which the conclusions are drawn, has been doubted. In 2010, Siegfriet wrote in Science News, ‘It is science’s dirtiest secret: the scientific method of testing hypothesis by a statistical analysis stands on a flimsy foundation. Statistical techniques for testing hypotheses ... have more flaws than Facebook’s private policies’. By convention, a ‘*p*’ value of less than 0.05 is declared statistically significant. But that was challenged by Wasserstein and Lazar in a statement released on behalf of the American Statistical Association [6]. It is not that the ‘*p*’ value should be discarded, but certainly the authors insist, instead of a binary less than or more than 0.05, we should quote the actual ‘*p*’ value.

In fact, the entire March 2019 issue of American Statistician is dedicated to ‘Statistical inference in the 21st century: A world beyond $p < 0.05$ ’. With more than 40 papers, the gist is that the ‘*p*’ value should be quoted in absolute numbers and not in a binary fashion, and even ‘*p*’ values outside the range may have clinical significance, albeit of a lesser degree. The issue takes the ‘dichotomania to the confidence intervals too’ and suggests a new term—‘compatibility intervals’. It is not that the values lying outside the interval are incompatible, but they are just a little less compatible than the ones within the interval [2]. ‘How do statistics so often lead scientists to deny differences that those not educated in statistics can plainly see? Lets be clear about what must stop: We should never conclude, there is ‘no difference’ or ‘no association’ just because a ‘*p*’ value is larger than a threshold set as 0.05 or, equivalently because a confidence interval includes zero’ [2] (refer Fig. 1). Even the dogmatic subscription to the 95% limit to confidence interval has been decried. Thus ‘*p*’ values and confidence/compatibility intervals are a continuum, and everything should not be seen as black and white, but there are shades of grey, which too may have clinical

✉ Om Prakash Yadava
op_yadava@yahoo.com

¹ National Heart Institute, New Delhi, India

Fig. 1 Beware false conclusions
[2] (Reproduced with permission)

significance. However, the authors issue a caveat—just as these new measures are suggested to improve on data analysis and their relevance; they are not a panacea to all the ills, and they themselves will be prone to fallacious interpretation [2].

Should ‘peer reviews’ be ‘peer reviewed’?

And to bring up the rear, cry the headlines of a commentary, ‘To maintain trust in science, lose the peer review’, by Drs Mazer and Mandrola on Medscapes online forum ‘theheart.org’ (Feb 19, 2019). This cryptic comment drew from a Cochrane Systematic Review of 28 studies, which surmised, ‘little empirical evidence is available to support the use of editorial peer review as a mechanism to ensure the quality of biomedical research.’ The review further concluded, ‘At present, the absence of evidence on efficacy and effectiveness cannot be interpreted as evidence of their absence’ [3]. The manuscripts are quite often reviewed superficially and even at times, with a conflict of interest biasing the recommendation. To make matters worse, there is reticence of the editors and the journals to take action against the rogues. In fact, quite often, the well intended peer review process is even bypassed by the predatory journals.

All this just go to show that all established dogmas need to be challenged and relooked with an open mind. The ‘art’ and ‘science’ of practice of medicine must undergo continuous evolution, as, is in fact, the dictum of very existence—change is inevitable. Even the epistemology of Indian philosophies includes not just one, but multiple, methods of getting ‘Pramana’ (Proof). Besides ‘Pratyaksa’ (perception)—the foremost; ‘Anumana’ (inference),

‘Upamana’ (Comparison) and ‘Sabda’ (testimony of experts) are other valid means of obtaining knowledge. But these methods do not subscribe to the scrutiny of conventional statistical techniques. It is therefore important that common sense should rule the interpretation of all studies and trials, and one must not just blindly follow the ‘letter’, but also try and feel and appreciate the ‘spirit’, behind the findings, and their applicability to the society, that they are designed to serve.

References

1. Fanaroff AC, Califf RM, Windecker S, Smith SC, Lopes RD. Levels of evidence supporting American College of Cardiology/American Heart Association and European Society of Cardiology guidelines, 2008–2018. *JAMA*. 2019;321:1069–80.
2. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567:305–307.
3. Jefferson T, Rudin M, Brodny FS, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database Syst Rev*. 2007. <https://doi.org/10.1002/14651858.MR000016.pub3>.
4. Bonow RO, Braunwald E. The Evidence Supporting Cardiovascular Guidelines: Is There Evidence of Progress in the Last Decade? *JAMA*. 2019;321:1053–1054.
5. Yadava OP. First Indian practice guidelines in cardiac surgery ... never too late!. *Indian J Thorac Cardiovasc Surg*. 2019;35:S1–S2.
6. Wasserstein RL, Lazar NA. The ASA’s statement on ‘p’ values: context, process and purpose. *Am Stat*. 2016;70:129–33.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.