



The dilemma of surgical research between evidences and experience, impact factor and innovation

Luca Viganò^{1,4} · Antonio Giuliani² · Fulvio Calise³

Received: 11 March 2019 / Accepted: 14 March 2019 / Published online: 20 March 2019
© Italian Society of Surgery (SIC) 2019

Keywords Surgical research · Impact factor · Evidence-based medicine · Randomized trial · IDEAL collaboration · Innovation

What is the strongest evidence in surgery? The lack of evidences to guide practice. Several justifications have been advanced, including difficulty randomizing patients for surgical interventions, large benefits from innovative procedures that do not need scientific confirmation, and peculiarities of surgical patients that preclude applicability of any fixed rule. In this scientific anomaly, surgical education became studded with myths that inappropriately gained the rank of evidences [1]. We are overwhelmed with surgical papers: in 2018, more than 50,000 manuscripts were published, including 11,000 about liver surgery, 6000 about colorectal surgery, 5000 about gastric surgery, and so on. Unfortunately, the quantity is not synonymous with quality. In 2010, a large amount of money was spent for biomedical research (about US\$ 240 billion) and a vast number of papers was produced (about 3 million articles, of which about half are published by 6000 publishers in 25,000 journals), but 85% of studies was classified as “avoidable waste” [2]. We concur with the Altman conclusion which was “we need less research, better research, and research done for the right reasons” [3]. But the way out from this status is not obvious.

Which compass can guide us? The impact factor (IF) and the evidence-based medicine (EBM) have been strongly suggested as solutions. IF is the key to decode the list of more than 400 surgical journals registered in the SCIMAGO database (www.scimagojr.com). Provided that it is impossible to read every published study, the IF helps us identifying the most relevant ones. But some (several) cautionary notes are mandatory. First, the IF largely differs among journals of different medical specializations, for example, oncologic or gastroenterological journals have much higher values than surgical ones. This introduces a sort of ranking of clinical studies that does not sound completely logical to us. Second, the IF reflects the mean value of the articles published by a journal and does not necessarily apply to all papers in the given journal. Third, IF has become the glittering pin of scientific journals. Randy Schekman, the winner of the 2013 Nobel Prize for Medicine, strongly criticized this situation [4]. Nowadays, the IF is a tool to sell a brand (the journal) and risks to be pursued in place of scientific purposes. As Schekman stated, “a paper can become highly cited because it is good science or because it is eye-catching, provocative or wrong”. A major bias in paper selection has been introduced. Surgical research means innovation, but innovation takes time to be recognized. Easy-to-understand and confirmatory messages are much easier to publish than innovative ones and collect more citations in a short period. Paula Stephan classified papers as ‘non-novel’, ‘moderately novel’ and ‘highly novel’ and compared how they were cited [5]. Highly novel papers were more likely to be either highly cited or ignored and tend to be published in journals with lower IF. The topics that became big hits took time (> 3 years, more likely after 15 years) to be recognized. This phenomenon, depicted by the diffusion of innovations theory, affected even the most famous technological innovations, such as smartphones, but conflicts with IF logic that

✉ Luca Viganò
luca.vigano@hunimed.eu

¹ Division of Hepatobiliary and General Surgery, Department of Surgery, Humanitas Clinical and Research Center, IRCCS, Rozzano, MI, Italy

² Department of Transplantation, Unit of Hepatobiliary Surgery and Liver Transplant Center, ‘A. Cardarelli’ Hospital, Naples, Italy

³ Center for Hepatobiliary and Pancreatic Surgery, Pinetagrando Hospital, Castel Volturno, CE, Italy

⁴ Department of Biomedical Sciences, Humanitas University, Via A. Manzoni, 56, 20089 Rozzano, Milano, Italy

considers the short period (2 years). Finally, IF is consistently one of the criteria used to evaluate academic careers, even in Italy where the threshold values for having access to academic positions are progressively increasing. The “publish or perish” philosophy risks to privilege strategies to catch IF points, rather than guaranteeing high-quality and innovative researches.

On the other hand, EBM is the most important attempt to put order in the mess of clinical research. EBM is commonly associated with the pyramid of evidences that put on the top the experimental studies, namely the randomized trials, and the critical appraisal of studies, including meta-analyses, evidence-based guidelines and systematic reviews. However, EBM is not the panacea. The Centre for EBM Outcome Monitoring Project (COMPare) team demonstrated that results are incorrectly reported in almost 90% of trials published in the top five medicine journals (*New England Journal of Medicine*, *Lancet*, *JAMA*, *The BMJ*, and *Annals of Internal Medicine*) [6]. In most cases, some outcomes prespecified in the study design were missing or, vice versa, some non-prespecified results were added. Considering meta-analyses, they are powerful tools used to sum up data, but they rely on published papers. Publication bias largely affect the available literature and, consequently, meta-analyses, with positive results being much more published than negative ones and positive secondary outcomes much more highlighted than negative primary ones [7]. Even the reliability of guidelines is cast into doubt. Major limitations have been reported in terms of completeness of literature review, independency in evaluations, and stakeholders' involvement [8]. New rules have been established to protect EBM. The EQUATOR (Enhancing the QUALity and Transparency Of health Research) network promoted specific checklists to guarantee transparent and accurate reporting for any kind of study (www.equator-network.org), including the AGREE (Appraisal of Guidelines for REsearch & Evaluation) recommendations for guidelines. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) (www.gradeworkinggroup.org) codified a rigorous approach to guidelines elaboration. The literature is systematically reviewed and evaluated, having the possibility to grade down evidences in case of bias, inconsistency, indirectness or imprecision.

Surgeons usually sit on the lowest rungs of the pyramid of evidences. In 2003, only 3% of publications in leading surgical journals were randomized trials [9]. This number had increased later on [10], but reluctance persists. Surgeons counterpose to the official EBM their “experience-based” medicine (eBM). Randomized trials are difficult or sometimes even impossible to realize because of high costs, large number of patients to collect, or ethical problems. Surrogate methods of analysis have therefore become largely popular, one for all the propensity score models, but they have

been associated with some major pitfalls [11, 12]. Further, EBM is considered inadequate to recognize peculiarities of patients and does not confer an appropriate dignity to innovation. How to heal this fracture? We should recover the original definition of EBM, in which clinical expertise, best evidences and patient values were combined together. Evidences are the key to define guidelines for large groups of people, but have to be merged with a case-by-case judgment and eBM in single patient management. The IDEAL collaboration did a hard job in this sense and gave probably the most relevant help to surgical research [10]. It described the stages of innovation in surgery (idea, development, exploration, assessment, long-term study) and set recommendations for stage-specific optimal research designs and outcomes to evaluate the new treatments (<http://www.ideal-collaboration.net/>). The IDEAL collaboration advanced concrete proposals to solve an impasse in surgical research. Case series studies should be replaced by prospective ones that can provide the basis for pre-trial evaluation. When randomized trials are not feasible, alternative prospective designs have been proposed, such as interrupted time series analysis that relies on a before–after comparison within a single population, rather than a comparison with a control group. It is up to surgeons to get involved in these proposals and reach the so craved position of high-evidence researchers.

In conclusion, surgical research needs more quality and good rules to drive its development. An evidence-based surgery is possible without limiting innovative impulses and without denying its peculiarities. Surgeons have to work in this direction. Updates in Surgery has the great opportunity to contribute to this process and guarantee high-quality research certified by the upcoming IF.

Compliance with ethical standards

Conflict of interest The authors have no source of funding and no conflict of interest to declare.

Ethical approval The study was approved by the local ethical committee.

Research involving human participants and/or animals The present study does not involve human participants and/or animals.

Informed consent Informed consent was obtained from each included patient.

References

1. Schein M (2004) Common myths in surgery. *Surg Rounds* 1:34–38
2. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, Al-Shahi Salman R, Chan AW, Glasziou P (2014)

- Biomedical research: increasing value, reducing waste. *Lancet* 383:101–104
3. Altman DG (1994) The scandal of poor medical research. *BMJ* 29:283–284
 4. Schekman R (2013) How journals like nature, cell and science are damaging science. <https://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science>. Accessed 9 Dec 2013
 5. Stephan P, Veugelers R, Wang J (2017) Reviewers are blinkered by bibliometrics. *Nature* 26(544):411–412
 6. Ioannidis JP, Caplan AL, Dal-Ré R (2017) Outcome reporting bias in clinical trials: why monitoring matters. *BMJ* 356:j408
 7. Every-Palmer S, Howick J (2014) How evidence-based medicine is failing due to biased trials and selective publication. *J Eval Clin Pract* 20:908–914
 8. Gavriilidis P, Roberts KJ, Askari A, Sutcliffe RP, Huo TL, Liu PH, Hidalgo E, Compagnon P, Lim C, Azoulay D (2017) Evaluation of the current guidelines for resection of hepatocellular carcinoma using the Appraisal of Guidelines for Research and Evaluation II instrument. *J Hepatol* 67:991–998
 9. Wente MN, Seiler CM, Uhl W, Büchler MW (2003) Perspectives of evidence-based surgery. *Dig Surg* 20:263–269
 10. McCulloch P, Feinberg J, Philippou Y, Koliass A, Kehoe S, Lancaster G, Donovan J, Petrinic T, Agha R, Pennell C (2018) Progress in clinical research in surgery and IDEAL. *Lancet* 392:88–94
 11. Hade EM, Lu B (2014) Bias associated with using the estimated propensity score as a regression covariate. *Stat Med* 33:74–87
 12. King G, Nielsen R (2018) “Why propensity scores should not be used for matching.” Political analysis. <https://gking.harvard.edu/files/gking/files/psnot.pdf>. Accessed 10 Nov 2018

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