

The future of evidence-based medicine: is the frog still boiling?

David J.J. Muckart^{1,2}, Manu L.N.G. Malbrain^{3–5}

¹*Department of Surgery, Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa*

²*Level I Trauma Unit and Trauma Intensive Care Unit, Inkosi Albert Luthuli Central Hospital, Durban, South Africa*

³*Department of Intensive Care Medicine and High Care Burn Unit, Ziekenhuis Netwerk Antwerpen, ZNA Stuivenberg Hospital, Antwerp, Belgium*

⁴*Department of Intensive Care, University Hospital Brussels (UZB), Jette, Belgium*

⁵*Faculty of Medicine, Free University of Brussels (VUB), Brussels, Belgium*

Anaesthesiology Intensive Therapy 2017, vol. 49, no 5, 329–335

ANALOGY BETWEEN LAW AND MEDICINE

As stated previously, for more than 2,000 years personal experience, anecdotes, and bias have dictated medical practice [1]. This editorial will elaborate further in depth on a paper written in 2013, entitled “Evidence-based medicine: are we boiling the frog?” concluding that in employing an analogy to criminal law, even if the evidence is beyond a reasonable doubt, it is rarely unequivocal or indisputable; thus, evidence is not synonymous with truth [1]. The opinions of medical leaders of successive eras have held sway and have done so from the days of Hippocrates and Galen. Dupuytren caused untold harm by proclaiming that under no circumstances could a structure as insignificant as the appendix be responsible for any abdominal mischief [2]. The traditional hierarchical training structure perpetuated such dogma whereby the consultant’s word was law.

Evidence-based medicine is defined as ‘The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.’ However, as medicine is not an exact science as such, there may be many different ways to treat and cure our patients. Medicine, therefore, is an uncertain science [1]. Medical practitioners have followed their legal colleagues and nowadays scientific proof, based on the best available evidence, is required to substantiate current practice [1]. In law, decisions in civil cases are made on the basis of “more likely than not” where the balance of probability needs to be only greater than fifty percent. Criminal verdicts, which may incur severe penalties, require the evidence to be beyond reasonable doubt. Scientific proof must reach a higher level of cer-

tainty, as compared with the legal system [3]. This search for scientific proof has resulted in the construction of the evidence pyramid with the supposed best level of evidence provided by meta-analysis at the apex, the least reliable being personal opinions at the base (Fig. 1). In between these, in descending order of accuracy, lie systematic reviews, randomized controlled trials, cohort studies, case-control studies, case series and case reports.

As with legal argument, the evidence provided by each level has been contested, strong opinions being voiced by opposing camps [4, 5]. Paradoxically, personal eye-witness testimony on which a verdict may hinge in a court of law is relegated to the least reliable level of evidence in medicine, namely the individual case report. Strong evidence in clinical research is reportedly underpinned by the randomized controlled trial (RCT) with appropriate numbers of patients, assignment, and blinding. From the compilation of all of these, a meta-analysis is the advocate of the standard of care. Adherence to the requirements for reliable clinical research will ensure, in the majority of studies, that the results are valid [5]. Even when performed within such strict confines, however, bias may confound the results. This is especially true of interventional trials, the studies to which the RCT is most pertinent.

NEGATIVE STUDIES INDUCE PUBLICATION BIAS

A meta-analysis is only as good as the studies selected and inevitably can include only those published. There exists, however, a publishing bias against studies with negative or inconclusive findings in both drug and equipment trials

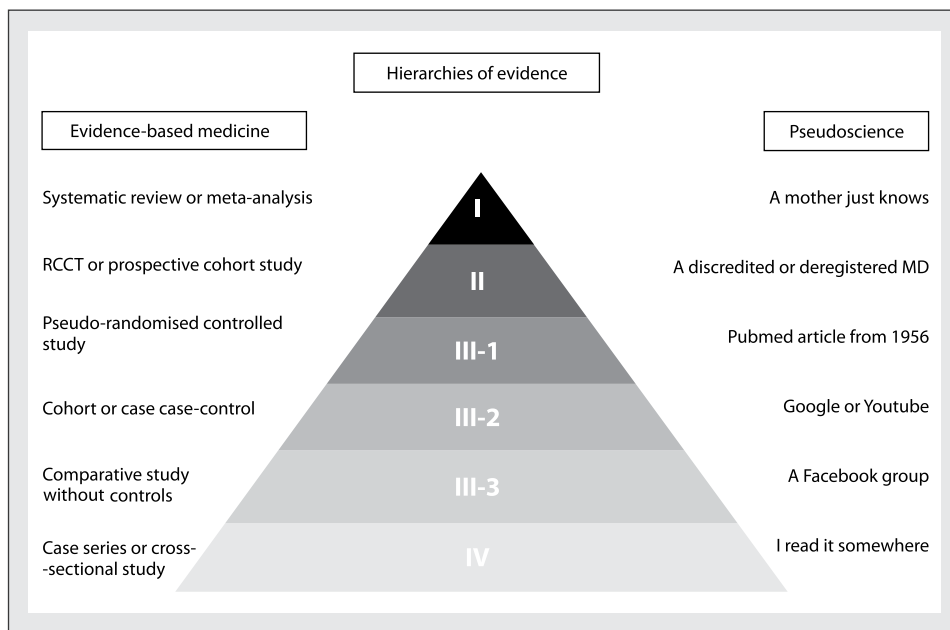


Figure 1. The hierarchy of evidence pyramid

[6–8]. Many reasons account for this. Investigators may lose interest in an inconclusive research result; sponsors may be unwilling to release detrimental information concerning their product. Personal opinion may sway reviewers and editors. The exclusion of unpublished data will inevitably skew the findings of any meta-analysis [9]. Clinical trials in which the results show a significant difference are three times more likely to be published than those with insignificant findings. The time to publication is also significantly longer with 80% of positive trials but only 20% of negative trials being published within ten years of completion [6]. An undetected positive bias may also exist in the form of duplicate publication [7]. The same study portraying a different aspect of the data may be published under an alternative sequence of authorship and be included twice in a meta-analysis. In the case of highly significant findings, the treatment effects will be overestimated.

For an unbiased meta-analysis, studies must be carefully scrutinized for potential duplication and a liberal time frame must be allowed for the inclusion of studies containing inconclusive or negative data. Unpublished information is also relevant, and for more than two decades a plea has been made for the registration of all interventional studies, thus allowing access to such data [8–11]. Failure to publish, or allow access to all information has been suggested as being tantamount to scientific misconduct [12]. Recently, the International Committee of Medical Journal Editors (ICMJE) has become insistent on the criteria for acceptance for publication, a move which appears to have succeeded in encouraging the registration of clinical trials [13].

Bias against negative findings goes beyond non-publication where results which are detrimental to the tested product may be deliberately suppressed, the positive aspects alone being published [14–18]. The last 20 years have seen a trend from a time when investigators controlled all aspects of a study from design, data collection, analysis, and publication to an era where the sponsor assumes control and offers financial incentives to enrol patients [11]. A contract between researchers and funders is often agreed upon whereby commercial sponsors control the database, perform the statistical analysis, and after that, present the finished article to the researchers. Unless the researchers have direct access to the entire data and are allowed to conduct an independent review, there remains the opportunity for sponsors to withhold damaging information or skew the data in favour of a positive outcome [18]. Information describing the detrimental effects may only surface at a later date after patient harm has occurred [19].

EFFECT OF SPONSORSHIP BY FOR-PROFIT ORGANIZATIONS

Funding of new drug trials is almost invariably, and understandably, provided by a parent company. Fiscal contributions go beyond the costs of delivering the drug and conducting the study, however, and financial incentives are often offered for individual researchers, either personally or for the department to which they belong. In a poll of 459 academic departments, almost two-thirds revealed financial ties with the pharmaceutical or medical equipment industry including acting in a consultative capacity,

on advisory boards, and speakers' bureaus [20]. Does this have an impact on study outcomes? The answer would appear to be that it does. An analysis of 159 trials involving 12 different specialties concluded that there was a significant finding in favour of the trial drug if the study was funded by for-profit organizations. This could not be explained by methodology, statistical analysis, or the type of study [21]. A similar review of meta-analyses found that 51% of studies funded by for-profit organizations found in favour of the experimental drug compared to only 16% of studies sponsored by non-profit organizations [22]. Trials funded by for-profit organizations were more than 5 times as likely to recommend the trial drug as the treatment of choice. This must cast doubt upon the validity of certain conclusions. As stated by Angell, a former editor of *the New England Journal of Medicine* for two decades, "Physicians can no longer rely upon the medical literature for valid and reliable information". She reluctantly concludes that these products are not nearly as effective as the publications suggest [17].

This is not a subjective appraisal and has been substantiated in the literature. An analysis of highly cited trials published in the three journals with the highest impact factor (*The New England Journal of Medicine*, *The Lancet*, *Journal of the American Medical Association*) and those with an impact factor greater than seven, showed that 30% of trials initially reporting highly significant positive findings were found in subsequent studies to either overestimate treatment effects or show no benefit [23], a phenomenon confirmed in more recent studies [24]. Admissible studies covered a broad cross-section of medical and surgical specialties. The danger lies in such initial studies being adopted blindly into clinical practice without confirmation which, due to the time-consuming nature of clinical research, may take years to produce. Scepticism is essential, especially when the initial results show an exceptional outcome. Unfortunately, it is these very studies which are more readily accepted without validation.

The effect of funding affects more than pharmacology or new device trials. Consensus guidelines and statements written by panels of experts are frequently supported by the industry, and the sponsoring company may have financial affiliations with the members of such panels [17, 25]. This has led to specific journals refusing to accept for publication consensus statements, reviews, or commentaries which recommend drugs, equipment or other profitable interventions if authors have declared a financial conflict of interest [25].

GHOST AUTHORS, GUEST AUTHORS, AND DATA FABRICATION

Ghost authorship takes two forms. In its benign form, professional medical writers skilled in the use of language may improve a manuscript without altering the scientific

content (perhaps they should be regarded as the real "spin doctors"). They may be acknowledged in the text but will not appear on the authors' list. A more malignant tendency has spread in industry-sponsored studies whereby the initial draft of the manuscript, namely that which shapes the final article, is compiled by company employees. Key opinion leaders that are academically affiliated, are often sourced as principal or second authors without having provided any substantial contribution to the study or article [26]. Articles in high-impact journals under the authorship of respected academics undoubtedly influence clinical practice. Representatives of the companies use them to promote their product while clinicians independently garner evidence from the current literature regarding best practice. In the most adverse scenario, pharmaceutical companies engage the services of contract research organizations to organize the trial, the company then analyses the data, use their employees to write the manuscript, source a respected academic to act as the principal author, and pay communication companies to accelerate publication [27]. The exact incidence of ghost authorship is unknown.

From painted mice to post-operative pain relief, colloids to cardiac protection, instances of trial misconduct and data fabrication have raised their ugly heads. This literary hydra is a cause for serious concern and casts an uncomfortable shadow over medical evidence. A recent analysis reported that 2% of scientists admitted to fabricating or modifying data at least once and one-third confessed to questionable research practices. Interrogation of colleagues revealed more alarming figures of 14% concerning data falsification, and 72% for dubious scientific behaviour [28].

STATISTICAL VERSUS CLINICAL SIGNIFICANCE

Statistical significance is the cornerstone of the bridge between clinical trials comparing two or more groups and their conclusions. Simply put, it produces a mathematical probability of whether the results of a study are due to chance, a five percent risk of the effects being falsely positive deemed acceptable. As amusingly described by Hall [29], this is not due to divine intervention but was the learned opinion of the statistician Fisher. In essence, this is, therefore, based on subjective expert opinion, the antithesis of evidence-based medicine.

However, in reality, clinical relevance has little or nothing to do with statistical significance and must not be confused with biological importance [29]. In fact, Luus *et al.* [30] suggest that clinically relevant differences and statistical significance concur only by coincidence. These authors pointed out that clinicians must have a good understanding of the results (and the statistical methods used), while statisticians need to understand the clinical problem in order to generate analytical results in line with meaningful clinical conclusions.

As the aim of trials is to determine whether the aspect under scrutiny will have an impact on clinical practice, the results should, therefore, be expressed in clinical and not mathematical terms. The latter are largely susceptible to sample size and clinical differences may be statistically significant, albeit meaningless [1]. The reverse also holds true if the sample size is too small. In his book “The Last Well Person”, Hadler [31] argues that an absolute reduction in the mortality rate of 1.9% in the West of Scotland statin trial [32], although statistically significant, is clinically irrelevant. The data however are expressed as a relative risk reduction of 29%, a more impressive statistic. He suggests that no study can control for all confounders and an absolute difference of less than 2% should be viewed with caution.

More recently, although first described almost 50 years ago, the concept of intention to treat (ITT) analysis for RCT's has gained widespread advocacy, especially amongst statistical purists [33]. Using this approach, data are analysed according to randomization regardless of dropouts, the crossover between treatment groups, or missing data. They argue that this technique assures homogeneity between randomized groups and minimizes a type I error (the incorrect rejection of the null hypothesis). Clinicians argue that this concept is flawed, increases the risk of a type II error (the false acceptance of the null hypothesis) and propose that two other techniques provide more meaningful clinical information. These are a per protocol (PP) analysis where only compliant patients are included, or in the event of crossovers an as-treated (AT) analysis where subjects who received a particular treatment are included regardless of the group to which they were initially assigned. If the number of patients who drop out or crossover is minimal, the ITT concept holds true, but if higher than 20% the results will be inaccurate [34]. Furthermore, the choice depends on the question to be answered. If this relates to the efficacy of an intervention, then PP or AT analyses are superior. This concept is ably demonstrated by the recent EVAR trial [35] where subjects deemed too unfit for open repair were randomly allocated to either EVAR or no intervention. In the latter group, 34% underwent repair for a variety of reasons. ITT analysis showed a significant reduction in aneurysm-related deaths but no improvement in overall survival at eight years. PP analysis not only showed a much more significant reduction in aneurysm-related fatalities but a marked increase in overall survival. Both effects were even more evident using the AT technique. Hall suggests that in addition to an ITT analysis, authors should declare non-compliance and provide a summary that excludes these patients [29].

The number needed to treat (NNT) is of even greater clinical relevance, the reciprocal of absolute risk reduction, which defines how many patients need to be treated for one to gain a benefit [1, 36], which may also be expressed

per 1,000 patients. Of importance, this calculation does not correlate with probability values but assesses clinical impact. Equally, or of perhaps greater concern is the number need to harm (NNH) which assesses the possible adverse consequences of a particular intervention. The POISE (Post Operative Ischaemic Evaluation) study epitomizes these concepts [37]. This is the largest randomized controlled trial to assess whether peri-operative beta blockade can lower the risks of post-operative cardiovascular events. A highly significant reduction of non-fatal myocardial infarctions was found in the treated group with an NNT of 66. The incidence of stroke, however, doubled with the NNH being 200. For every three patients spared a cardiac event, one would potentially suffer a cerebral insult. In those who suffered a stroke in the treatment group, however, only 15% regained full function while 26% were left severely incapacitated. The choice between the risk and sequelae of a non-fatal myocardial infarct versus a disabling stroke is a matter of clinical judgement and patient preference, not mathematical probability [1].

THE EVIDENCE PYRAMID

Meta-analysis lies at the summit of the evidence pyramid and is regarded as the final arbiter in areas of uncertainty with conflicting or scarce data (Fig. 1). In essence, it is a statistical averaging technique undertaken in the hope that by pooling data, any errors in individual study design or implementation will be minimized. Errors in clinical trials may be random or systematic [38]. The former is unpredictable and may skew data in both positive and negative directions. An increase in sample size reduces its occurrence [1]. Systematic errors are not eliminated by increasing the sample size, and arise when a trend in the data occurs that is actually false [1]. This occurs as a result of three types of bias, namely selection, misclassification, and confounding.

SELECTION BIAS

Selection bias occurs when a test is inadvertently skewed to favour a subset of patients, as in that cited by Tobin and Jubran [38] where patients who failed a weaning test were automatically excluded from a trial of spontaneous breathing.

MISCLASSIFICATION BIAS

Misclassification bias describes the error of placing patients in an incorrect category, resulting in a heterogeneous rather than homogeneous population being placed under scrutiny [1]. This is especially true where standard therapies are normally titrated against physiological end-points rather than fixed dose regimens [1, 39]. Deans *et al.* [39] cite the ARDS low tidal volume trial as a prime example where patients were randomized to fixed tidal volumes of either 6 mL kg⁻¹ or 12 mL kg⁻¹ whereas the standard practice would be to titrate

treatment in accordance with airway pressures and compliance. Randomisation to a fixed dose would result in some patients in the low tidal volume group being underventilated and some in the high volume group being overventilated. As emphasized by Vincent [40], the identical scenario pertains to transfusion triggers. Younger patients without coronary artery disease may tolerate a lower haemoglobin level than the elderly cardiopath and conversely, overtransfusion in the young may lead to a detrimental effect [1, 40]. Such insufficient or excessive treatment may contribute substantially to differences in the trial results. Thus, ignoring heterogeneity violates sound statistical practice [41]. In truth, there is no such entity as a perfectly matched RCT, for although the control and treatment groups may demonstrate no statistically significant difference between the means of any parameter, each group, even if normally distributed, will inevitably be heterogeneous.

CONFOUNDING BIAS

Confounding bias refers to the mistaken relationship found between two variables because of a third unaccounted factor. Even if selected studies are homogeneous, disagreement may arise amongst reviewers who, when presented with identical sets of information, arrive at diametrically opposite conclusions [42]. As with any other scientific publication, meta-analyses must be interpreted with caution. For over fifteen years, there has been an appeal for individual patient data to be used, but even if sourced and included, systematic error, publication, and selection bias may still exist [43]. The most recent example is that of the systematic review and meta-analysis of hydroxyethyl starch for volume resuscitation which purports to find a higher incidence of acute kidney injury with the use of starch solutions [44]. Selected for inclusion in the analysis were such diverse groups as burns, trauma, septic and non-septic ICU patients, organ donors, acute pancreatitis and those resuscitated post cardiac arrest. This study concludes by stating, "Clinical use of hydroxyethyl starch for acute volume resuscitation is not warranted due to serious safety concerns", a conclusion that may pertain to a sub-group somewhere but cannot be applied to such a diverse patient population. The consequence of such unfounded statements in a high impact publication is that many clinicians will abandon the use of colloids and revert to crystalloids which are known to have serious adverse effects if used as the sole resuscitation fluid.

The precept that systematic reviews should exclude observational studies has recently been challenged, and in situations where randomized trials are either few or of

inferior design, non-randomised trials should be eligible for meta-analysis [45]. Given the multitude of problems in conducting randomized controlled trials in the critically ill, Vincent [40] has suggested that these should be abandoned and more attention paid to observational studies. This, he argues, would allow enrolment of almost the entire patient population under investigation with fewer exclusions, making such an approach more clinically relevant.

ARE WE STILL BOILING THE FROG?

But what of the boiling frog and its relevance to evidence-based medicine? There is a physiological anecdote (although this may be an urban legend¹) that if a frog is placed in boiling water, it will understandably leap out immediately. If the water is initially tepid however, and slowly heated to boiling point, the frog will remain until boiled alive [1]. This example has been used in various scenarios including economics and global warming, to illustrate the concept that slow change may pass unrecognized until harm occurs [1]. The current United Nations Congress on climate change and international fiscal crisis epitomize this concept. Numerous examples exist in medical practice across all specialties; the pulmonary artery catheter in critical care; in surgery abdominal tension sutures when the abdomen is difficult or painful to close; bowel preparation for colonic surgery; the list could be considerable. Proof of benefit of any intervention is an understandable and laudable aim of any clinician and this, coupled with the desire to do no harm, has been the driving force for evidence-based medicine. This concept is not scientifically perfect and must not be viewed as exclusive [45, 46], especially in relation to practice guidelines. Despite claims that these represent the best standard of care, this does not hold true for all patients with the same disease. As observed by Osler, "Variability is the law of life and no two individuals react alike and behave alike under the abnormal conditions which we know as disease. The good physician treats the disease, the great one treats the patient" [47]. Evidence-based guidelines not only ignore patient heterogeneity but depending on the quality of evidence, may not define optimal treatment [48, 49]. Blind acceptance may cause harm, as illustrated by the use of potassium-sparing diuretics for ventricular dysfunction [50]. The recommendation for spironolactone as the diuretic of choice by the American Heart Association resulted in a more than four-fold increase in prescriptions. Over the ensuing three years, there was a four-fold increase in hospital admissions for hyperkalaemia and an almost seven-fold increase in death from this complication.

¹The story was based on nineteenth-century experiments in which frogs were shown to stay in heating water as long as it was heated very slowly. The validity of the experiments is, however, disputed. Professor Douglas Melton, Harvard University Biology Department, says: "If you put a frog in boiling water, it won't jump out. It will die. If you put it in cold water, it will jump before it gets hot — they don't sit still for you". (See also: <https://www.fastcompany.com/26455/next-time-what-say-we-boil-consultant>)

Although whimsically sarcastic, Smith and Pell rightfully conclude there is no evidence that a parachute prevents major harm when jumping from an aircraft, but common sense should prevail. They suggest the same should apply when considering the potential risk and benefit of interventions [51]. In light of the above it is essential that evidence-based medicine be viewed with an open mind. As in criminal law, even if the evidence is beyond reasonable doubt, it is rarely unequivocal or indisputable; evidence is not synonymous with truth [1]. From initially tepid waters the zeal for evidence-based practice has now reached boiling point and if the shortcomings are not appreciated, evidence-based medicine may itself become a boiled frog. Even in modern practice Osler's aphorism still holds true, namely "Medicine is a science of uncertainty and an art of probability".

ACKNOWLEDGEMENTS

Parts of this editorial were published previously under the Open Access CC BY Licence 4.0 [1].

David Muckart is Associate Professor of Surgery at the Nelson R Mandela School of Medicine, University of Kwa-Zulu-Natal, Durban, South Africa, and Chief Specialist at the Level I Trauma Unit and Trauma Intensive Care Unit, Inkosi Albert Luthuli Central Hospital, Durban. His clinical and research activities concentrate on the management of the critically injured, and he has a keen interest in the history of medicine and surgery.

Manu Malbrain is Professor of Medicine at the Free University of Brussels and ICU Director at the University Hospital in Brussels, Belgium. He is founding President of WSACS (The Abdominal Compartment Society) and current Treasurer, he is also member of the medical advisory Board of Pulsion Medical Systems (part of Maquet Getinge group) and consults for ConvaTec, Acelity, Spiegelberg and Holtech Medical. He is also co-founder of the International Fluid Academy (IFA). This article is endorsed by the IFA. The mission statement of the IFA is to foster education, promote research on fluid management and haemodynamic monitoring, and thereby improve the survival of the critically ill by bringing together physicians, nurses, and others from throughout the world and from a variety of clinical disciplines. The IFA is integrated within the not-for-profit charitable organization iMERiT, International Medical Education and Research Initiative, under Belgian law. The IFA website (<http://www.fluidacademy.org>) is now an official SMACC affiliated site (Social Media and Critical Care) and its content is based on the philosophy of FOAM (Free Open Access Medical education — #FOAMed). The site recently received the HONcode quality label for medical education (<https://www.healthonnet.org/HONcode/Conduct.html?HONConduct519739>).

References:

- Muckart DJJ. Evidence-based medicine - are we boiling the frog? *S Afr Med J*. 2013; 103(7): 447–448, indexed in Pubmed: [23802201](#).
- Zimmerman L, Cartwright F. The Development of Modern Surgery. *The American Historical Review*. 1968; 74(1): 106, doi: [10.2307/1857639](#).
- Miller DW, Miller CG. On Evidence, Medical and Legal. *J American Physicians and Surgeons*. 2005; 10: 70–75.
- Karanicolas PJ, Kunz R, Guyatt GH. Point: evidence-based medicine has a sound scientific base. *Chest*. 2008; 133(5): 1067–1071, doi: [10.1378/chest.08-0068](#), indexed in Pubmed: [18460513](#).
- Tobin MJ. Counterpoint: evidence-based medicine lacks a sound scientific base. *Chest*. 2008; 133(5): 1071–4; discussion 1074, doi: [10.1378/chest.08-0077](#), indexed in Pubmed: [18460514](#).
- Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ*. 1997; 315(7109): 640–645, indexed in Pubmed: [9310565](#).
- Montori VM, Smieja M, Guyatt GH. Publication bias: a brief review for clinicians. *Mayo Clin Proc*. 2000; 75(12): 1284–1288, doi: [10.4065/75.12.1284](#), indexed in Pubmed: [11126838](#).
- Gregor S, Maegele M, Sauerland S, et al. Negative pressure wound therapy: a vacuum of evidence? *Arch Surg*. 2008; 143(2): 189–196, doi: [10.1001/archsurg.2007.54](#), indexed in Pubmed: [18283145](#).
- Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet*. 2004; 363(9418): 1341–1345, doi: [10.1016/S0140-6736\(04\)16043-1](#), indexed in Pubmed: [15110490](#).
- Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol*. 1986; 4(10): 1529–1541, doi: [10.1200/JCO.1986.4.10.1529](#), indexed in Pubmed: [3760920](#).
- DeAngelis CD, Fontanarosa PB. Impugning the integrity of medical science: the adverse effects of industry influence. *JAMA*. 2008; 299(15): 1833–1835, doi: [10.1001/jama.299.15.1833](#), indexed in Pubmed: [18413880](#).
- Chalmers I. Underreporting research is scientific misconduct. *JAMA*. 1990; 263(10): 1405–1408, indexed in Pubmed: [2304220](#).
- Laine C, Horton R, De Angelis C, et al. International Committee of Medical Journal Editors (ICMJE), International Committee of Medical Journal Editors (ICMJE). Clinical trial registration – looking back and moving ahead. *N Engl J Med*. 2007; 356: 2734–2736.
- Psaty BM, Kronmal RA. Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment: a case study based on documents from rofecoxib litigation. *JAMA*. 2008; 299(15): 1813–1817, doi: [10.1001/jama.299.15.1813](#), indexed in Pubmed: [18413875](#).
- Melander H, Ahlqvist-Rastad J, Meijer G, et al. Evidence based medicine-selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ*. 2003; 326(7400): 1171–1173, doi: [10.1136/bmj.326.7400.1171](#), indexed in Pubmed: [12775615](#).
- Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008; 358(3): 252–260, doi: [10.1056/NEJMsa065779](#), indexed in Pubmed: [18199864](#).
- Angell M. Industry-sponsored clinical research: a broken system. *JAMA*. 2008; 300(9): 1069–1071, doi: [10.1001/jama.300.9.1069](#), indexed in Pubmed: [18768418](#).
- Steinbrook R. Gag clauses in clinical-trial agreements. *N Engl J Med*. 2005; 352(21): 2160–2162, doi: [10.1056/NEJMp048353](#), indexed in Pubmed: [15917381](#).
- Rennie D. Trial registration: a great idea switches from ignored to irresistible. *JAMA*. 2004; 292(11): 1359–1362, doi: [10.1001/jama.292.11.1359](#), indexed in Pubmed: [15355937](#).
- Campbell EG, Weissman JS, Ehringhaus S, et al. Institutional academic industry relationships. *JAMA*. 2007; 298(15): 1779–1786, doi: [10.1001/jama.298.15.1779](#), indexed in Pubmed: [17940234](#).
- Kjaergard LL, Als-Nielsen B. Association between competing interests and authors' conclusions: epidemiological study of randomised clinical trials published in the *BMJ*. *BMJ*. 2002; 325(7358): 249, indexed in Pubmed: [12153921](#).
- Als-Nielsen B, Chen W, Glud C, et al. Association of funding and conclusions in randomized drug trials. *JAMA*. 2003; 290(7): 921, doi: [10.1001/jama.290.7.921](#).
- Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005; 294(2): 218–228, doi: [10.1001/jama.294.2.218](#), indexed in Pubmed: [16014596](#).

24. Pereira TV, Horwitz RL, Ioannidis JPA. Empirical evaluation of very large treatment effects of medical interventions. *JAMA*. 2012; 308(16): 1676–1684, doi: [10.1001/jama.2012.13444](https://doi.org/10.1001/jama.2012.13444), indexed in Pubmed: [23093165](https://pubmed.ncbi.nlm.nih.gov/23093165/).
25. Conflicts of interests and investments. *CMAJ*. 2004; 171(11): 1313, indexed in Pubmed: [16523541](https://pubmed.ncbi.nlm.nih.gov/16523541/).
26. Ross JS, Hill KP, Egilman DS, et al. Guest authorship and ghostwriting in publications related to rofecoxib: a case study of industry documents from rofecoxib litigation. *JAMA*. 2008; 299(15): 1800–1812, doi: [10.1001/jama.299.15.1800](https://doi.org/10.1001/jama.299.15.1800), indexed in Pubmed: [18413874](https://pubmed.ncbi.nlm.nih.gov/18413874/).
27. Sismondo S. Ghost management: how much of the medical literature is shaped behind the scenes by the pharmaceutical industry? *PLoS Med*. 2007; 4(9): e286, doi: [10.1371/journal.pmed.0040286](https://doi.org/10.1371/journal.pmed.0040286), indexed in Pubmed: [17896859](https://pubmed.ncbi.nlm.nih.gov/17896859/).
28. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One*. 2009; 4(5): e5738, doi: [10.1371/journal.pone.0005738](https://doi.org/10.1371/journal.pone.0005738), indexed in Pubmed: [19478950](https://pubmed.ncbi.nlm.nih.gov/19478950/).
29. Hall J. How to dissect surgical journals: VIII — Comparing outcomes. *ANZ Journal of Surgery*. 2011; 81(3): 190–196, doi: [10.1111/j.1445-2197.2010.05358.x](https://doi.org/10.1111/j.1445-2197.2010.05358.x).
30. Luus HG, Muller FO, Meyer BH. Statistical significance versus clinical relevance. *S Afr J Med*. 1989; 76(568): 570.
31. Hadler NM. *The Last Well Person*. McGill-Queen's University Press, Montreal and Kingston. 2004: 35–43.
32. Shepherd J, Cobbe SM, Ford I, et al. West of Scotland Coronary Prevention Study Group. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med*. 1995; 333(20): 1301–1307, doi: [10.1056/NEJM199511163332001](https://doi.org/10.1056/NEJM199511163332001), indexed in Pubmed: [7566020](https://pubmed.ncbi.nlm.nih.gov/7566020/).
33. Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol*. 1992; 21(5): 837–841, indexed in Pubmed: [1468842](https://pubmed.ncbi.nlm.nih.gov/1468842/).
34. Armijo-Olivo S, Warren S, Magee D. Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: a review. *Physical Therapy Reviews*. 2013; 14(1): 36–49, doi: [10.1179/174328809x405928](https://doi.org/10.1179/174328809x405928).
35. Sweeting MJ, Patel R, Powell JT, et al. EVAR Trial Investigators, United Kingdom EVAR Trial Investigators. Endovascular repair of aortic aneurysms in patients physically ineligible for open repair. *N Engl J Med*. 2010; 362(20): 1872–1880, doi: [10.1056/NEJMoa0911056](https://doi.org/10.1056/NEJMoa0911056), indexed in Pubmed: [20382982](https://pubmed.ncbi.nlm.nih.gov/20382982/).
36. Davidson RA. Does it work or not: clinical versus statistical significance. *Chest*. 1994; 106: 932–934.
37. Devereaux PJ, Yang H, Yusuf S, et al. POISE Study Group. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet*. 2008; 371(9627): 1839–1847, doi: [10.1016/S0140-6736\(08\)60601-7](https://doi.org/10.1016/S0140-6736(08)60601-7), indexed in Pubmed: [18479744](https://pubmed.ncbi.nlm.nih.gov/18479744/).
38. Tobin MJ, Jubran A. Meta-analysis under the spotlight: focused on a meta-analysis of ventilator weaning. *Crit Care Med*. 2008; 36(1): 1–7, doi: [10.1097/01.CCM.0000297883.04634.11](https://doi.org/10.1097/01.CCM.0000297883.04634.11), indexed in Pubmed: [18007269](https://pubmed.ncbi.nlm.nih.gov/18007269/).
39. Deans KJ, Minneci PC, Danner RL, et al. Practice misalignments in randomized controlled trials: Identification, impact, and potential solutions. *Anesth Analg*. 2010; 111(2): 444–450, doi: [10.1213/ane.0b013e3181aa8903](https://doi.org/10.1213/ane.0b013e3181aa8903), indexed in Pubmed: [19820238](https://pubmed.ncbi.nlm.nih.gov/19820238/).
40. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med*. 2010; 38(10 Suppl): S534–S538, doi: [10.1097/CCM.0b013e3181f208ac](https://doi.org/10.1097/CCM.0b013e3181f208ac), indexed in Pubmed: [21164394](https://pubmed.ncbi.nlm.nih.gov/21164394/).
41. Greenland S. Can meta-analysis be salvaged? *Am J Epidemiol*. 1994; 140(9): 783–787, indexed in Pubmed: [7977288](https://pubmed.ncbi.nlm.nih.gov/7977288/).
42. Shrier I, Boivin JF, Platt RW, et al. The interpretation of systematic reviews with meta-analyses: an objective or subjective process? *BMC Med Inform Decis Mak*. 2008; 8: 19, doi: [10.1186/1472-6947-8-19](https://doi.org/10.1186/1472-6947-8-19), indexed in Pubmed: [18495019](https://pubmed.ncbi.nlm.nih.gov/18495019/).
43. Publication Bias in Meta-Analysis. 2006, doi: [10.1002/0470870168](https://doi.org/10.1002/0470870168).
44. Zarychanski R, Abou-Setta AM, Turgeon AF, et al. Association of hydroxyethyl starch administration with mortality and acute kidney injury in critically ill patients requiring volume resuscitation: a systematic review and meta-analysis. *JAMA*. 2013; 309(7): 678–688, doi: [10.1001/jama.2013.430](https://doi.org/10.1001/jama.2013.430), indexed in Pubmed: [23423413](https://pubmed.ncbi.nlm.nih.gov/23423413/).
45. Shrier I, Boivin JF, Steele RJ, et al. Should meta-analyses of interventions include observational studies in addition to randomized controlled trials? A critical examination of underlying principles. *Am J Epidemiol*. 2007; 166(10): 1203–1209, doi: [10.1093/aje/kwm189](https://doi.org/10.1093/aje/kwm189), indexed in Pubmed: [17712019](https://pubmed.ncbi.nlm.nih.gov/17712019/).
46. Holmes D, Murray SJ, Perron A, et al. Deconstructing the evidence-based discourse in health sciences: truth, power and fascism. *Int J Evid Based Healthc*. 2006; 4(3): 180–186, doi: [10.1111/j.1479-6988.2006.00041.x](https://doi.org/10.1111/j.1479-6988.2006.00041.x), indexed in Pubmed: [21631765](https://pubmed.ncbi.nlm.nih.gov/21631765/).
47. Osler W. On the educational value of the medical society. *The Boston Medical and Surgical Journal*. 1903; 148(11): 275–279, doi: [10.1056/nejm190303121481101](https://doi.org/10.1056/nejm190303121481101).
48. Woolf S. Do clinical practice guidelines define good medical care? *Chest*. 1998; 113(3), doi: [10.1378/chest.113.3_supplement.166s](https://doi.org/10.1378/chest.113.3_supplement.166s).
49. Tobin M. Rebuttal From Dr. Tobin. *Chest*. 2008; 133(5): 1076–1077, doi: [10.1378/chest.08-0271](https://doi.org/10.1378/chest.08-0271).
50. Juurlink DN, Mamdani MM, Lee DS, et al. Rates of hyperkalemia after publication of the Randomized Aldactone Evaluation Study. *N Engl J Med*. 2004; 351(6): 543–551, doi: [10.1056/NEJMoa040135](https://doi.org/10.1056/NEJMoa040135), indexed in Pubmed: [15295047](https://pubmed.ncbi.nlm.nih.gov/15295047/).
51. Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ*. 2003; 327(7429): 1459–1461, doi: [10.1136/bmj.327.7429.1459](https://doi.org/10.1136/bmj.327.7429.1459), indexed in Pubmed: [14684649](https://pubmed.ncbi.nlm.nih.gov/14684649/).

Corresponding author:

Manu L.N.G. Malbrain

Department of Intensive Care Medicine, University Hospital Brussels (UZB), Jette, Belgium

Faculty of Medicine and Pharmacology, Free University Brussels

(VUB), Brussels, Belgium

e-mail: manu.malbrain@uzbrussel.be