

The laudable objective of ensuring that various facets of society are represented by insisting on having non-medical people in the IECs, and mandating that a lawyer or judge, social scientist or representative of a non-governmental voluntary agency, a philosopher/ethicist/theologian and a lay person from the community should all be on the IEC, seems not to be achieving the purpose of a broad consultation. It appears that, often, the non-medical persons on the committee are in awe of the medical persons and speak little. The idea of 'discourse ethics' as elucidated by Jurgen Habermas and Karl Otto-Apel seems to be absent. It appears that capacity building is needed here too. Developing and encouraging expertise in the field of ethics is an urgent necessity.

Monitoring

"Once IEC gives a certificate of approval it is the duty of the IEC to monitor the approved studies, therefore an oversight mechanism should be in place." (1: 19)

It is doubtful if many (or even any) IECs in the country are following this guideline. At least some members of the IEC would have to work full time on only the IEC in order to be able to effectively monitor trials, especially in large organisations like the AIIMS where a large number of trials are carried out.

Conclusion

The ICMR Ethical Guidelines for Biomedical Research on Human Participants is a strong starting point for the oversight and safe conduct of clinical trials in India. In order to translate intentions into practice, further development of the guidelines is required.

References

1. Indian Council of Medical Research. Ethical guidelines for biomedical research on human participants [Internet]. New Delhi: ICMR; 2006 [cited 2011 Aug 29]. 111 p. Available from: www.icmr.nic.in/ethical_guidelines.pdf
2. How ethical are for-profit institutional review boards? *Lancet*. 2009 Apr 25; 373(9673):1400.

Evidence-based medicine: can the evidence be trusted?

PRATHAP THARYAN

Professor of Psychiatry, Prof BV Moses & ICMR Centre for Advanced Research in Evidence-Informed Healthcare, Christian Medical College, Vellore 632002, Tamil Nadu INDIA e-mail: prathap@cmcvellore.ac.in

Abstract

Empirical research indicates that much of the evidence required for the practice of evidence-based medicine cannot be trusted. The research agenda has been hijacked by those with vested interests within industry and academia, determining what research is funded and how it is done and reported. Unnecessary, inappropriate, or poorly designed and reported research results in suboptimal health outcomes. Many well-reported randomized controlled trials are designed to deceive by their choice of comparators and outcomes, and manipulation of statistics to produce desired outcomes that are selectively reported. Undisclosed conflict of interest, ghost-writing, the manufacturing of disease to increase drug marketing, and the marketing of research disguised as education are common. Understanding the many ways in which research is used to deceive, rather than reliably inform health decisions, and reclaiming the research agenda, is the collective responsibility of the scientific community and civil society.

Evidence-Based Medicine (EBM) refers to the process of making medical decisions that are consistent with evidence from relevant research and envisages a therapeutic alliance between research-evidence, clinicians and patients (1). The linchpin of this alliance is the astute clinician who has the resources and skills to readily access and to critically appraise evidence from research. For the potential benefits of EBM to be fully realized in improved health outcomes for patients, many caveats apply; the most important one is whether the evidence from research can be trusted.

Can the evidence be trusted?

The randomized controlled trial (RCT) is considered the least biased study design to answer questions concerning the efficacy of interventions. However, the current research agenda is mostly set by those with vested interests. This leads to promising findings that are published, widely disseminated, and frequently cited, but are often not free from various biases and distortions of the "truth" (2). Health-outcomes are often unsatisfactory when this "evidence" is used in daily clinical practice. Wary clinicians are therefore resistant to subjugating their clinical experience to recommendations based on research.

This does not mean that one should discredit all research, or reject evidence-based approaches to informing healthcare decisions, entirely. On the contrary, if EBM is viewed as a continuously evolving heuristic structure for optimising clinical practice, this provides opportunities to redefine our position on what we consider evidence that we can trust, and to help re-shape the research agenda.

What determines research evidence that is valid and ethical?

For research to be scientifically valid and ethical, it must be relevant to the populations it is conducted in, and be conducted with the purpose of advancing science and reducing uncertainties in clinical care. It must use methods that ensure reliable results (internal validity), by being free from bias (systematic errors in the design, conduct, reporting and interpretation) and confounding (prognostic variables in participants that independently produce the outcome of interest rather than the intervention), and the effects of chance (random errors). It must also produce generalisable results (external validity) that are relevant to other populations. It must be accountable and be approved by a body of one's peers (ethics and research review, editorial and peer review); be transparent in its methods; and be fully and accurately reported. It must also be participatory in that the outcomes used should be relevant (in terms of what these outcomes are, and the magnitude of purported benefits) to the people who use the results of the research (patients, their carers, and health care providers). This requires collaborative, altruistic, and well-informed partnerships in designing, conducting and disseminating research evidence.

Why can't much of the evidence be trusted?

1. The hijacked research agenda

The motives for conducting research are often determined by considerations other than the advancement of science or the promotion of better health outcomes. Many research studies are driven by the pressure to obtain post-graduate qualifications, earn promotions, obtain tenured positions, or additional research funding; many others are conducted for financial motives that benefit shareholders, or lead to lucrative patents.

The majority of clinical trials conducted world-wide are done to obtain regulatory approval and a marketing licence for new drugs. These regulations often require only the demonstration of the superiority of a new drug over placebo and not over other active interventions in use. It is easier and cheaper to conduct these trials in countries with lower wages, lax regulatory requirements, and less than optimal capacity for ethical oversight. It is therefore not surprising that the focus of research does not reflect the actual burden of disease borne by people in the countries that contribute research participants, nor address the leading causes of the global burden of disease (5). Some "seeding" trials, conducted purportedly for the purpose of surveillance for adverse effects, are often only a ploy to ensure brand loyalty among participating clinician-researchers (6).

This hijacked research agenda perpetuates further research of a similar nature that draws more researchers into its lucrative embrace, entrenching the academic direction and position statements of scientific societies and academic associations. Funders and researchers are also deterred from pursuing more relevant research, since the enmeshed relationship between academic institutions and industry determines what research is funded (mostly drugs at the expense of other interventions), and even how research is reported; thus hijacking the research agenda even further away from the interests of science and society (7).

2. Science is cumulative but research is often spawned in splendid isolation

The biased research agenda often results in research studies that are designed and conducted without reference to the cumulative evidence from similar or identical studies performed previously. One survey of 1523 trials published from 1963 to 2004 estimated that less than 25% of previously reported RCTs were cited by subsequent RCTs (8). While research findings ought to be replicated in different settings and with sufficient numbers of people to provide generalisable and reliable results, the value of replicative research lies in learning from, and improving on, the methods and findings from previous research studies.

Examples abound of instances where interventions rushed into clinical practice based on insufficient or inappropriate evidence were proven to be ineffective, or even harmful, once the results of systematic reviews from good quality trials were made known (9). Other examples exist of the delayed recommendations of life-saving treatments by more than a decade, and the routine endorsement of ineffective, and even harmful treatments, in spite of available evidence that had been neglected (10). Ignoring what is already known before doing new research is unethical, as it can lead to wasted resources, and lead to unnecessary harms.

Even if new research is preceded by a search for relevant prior research, what is easily accessible in the form of published research represents only a fraction of what research is done and reported (4). The unwarranted belief in the efficacy of new interventions (optimism bias), and the tendency to uncritically accept evidence (often biased) that supports prior beliefs (confirmation bias) leads to biased conclusions confirming what one wants to believe in and rejecting findings that challenge one's beliefs.

3. Deception due to reliance on evidence from biased study designs

Every research question needs to be answered by a research design that offers the best possibility of providing reliable (unbiased) answers. Observational studies (cohort studies and case-control studies) offer the best design to answer questions regarding aetiology or causation of diseases, detection of rare events with interventions, or harms that take a long time to develop (11). While observational study designs may offer an alternative to RCTs when they are impractical, or unethical, such designs are limited by biases due to known and unknown confounders that large RCTs avoid by virtue of randomisation. Randomisation creates comparison groups that have similar prognoses at the outset, by balancing known and unknown confounders between the intervention arms (12). Reliance on observational data to inform decisions regarding interventions can be seriously misleading (9), and empirical evidence indicates that observational studies tend to over-estimate the beneficial effects of interventions compared to RCTs (13).

4. Deception due to the biased design of RCTs

For RCTs to provide results that are reliable, relevant and of use to clinicians and patients, the protocol must incorporate elements that minimise the effects of bias. They must have sufficiently large sample sizes to detect important differences in comparison groups, should they exist; use clinically meaningful comparison groups that will help resolve substantial uncertainty about which of the trial interventions would benefit a patient most; and analyse, and interpret results accurately and appropriately. They must also be reported without distortion of what was intended, done, and analysed (12). Ignorance of research methods, lack of regulatory or ethical oversight, and the pressures of the current research agenda contribute to research that does not fulfil these requirements.

Sample sizes and deceptive results

Many trials do not report calculations on which the sample size was estimated, often leading to sample sizes insufficient to detect even important differences between interventions (for primary, let alone secondary outcomes) (2). Those that do report such estimations often use inappropriate, or grossly over-optimistic, assumptions about expected effects, leading to small trials that lead to false negative results due to insufficient power to detect even important differences. They may also reach false positive conclusions due to biased conduct or reporting, or due to chance. A trial of an intervention that does not work may produce false positive results 5% of the time, by chance, at the conventional level of statistical significance of a P value <0.5. This 5% of trials is also more likely to be published than the negative or inconclusive results that will be seen with the other 95% of trials. These published trials are what clinicians read about, and that influence practice. Unless they are balanced by the negative trials, this leads to published evidence that ought not to be trusted but, sadly, often is.

Biases in the design, conduct and reporting of randomised trials

The RCT is considered the "gold standard" for evaluating the effects of interventions, but is also subject to biases that arise in design, conduct, interpretation and reporting. For the results of RCTs to be properly evaluated, transparent reporting is required of those dimensions of conduct that minimise biases that have been identified by empirical research as leading to erroneous results.

These domains are methods to prevent selection bias (randomisation and allocation concealment); performance bias (blinding of participants and study personnel), detection bias (blinding of outcome assessors for subjectively reported outcomes), attrition bias (accounting for all patients randomised in the analyses), and reporting biases: inadequate outcome reporting, and selective reporting of outcomes (referring to the trial protocol or trials registration document, if available). Empirical evidence indicates that inadequate methods to prevent or minimise the risk of bias, particularly poor allocation concealment, are associated with erroneous and unpredictable treatment effects, particularly when subjectively reported outcomes are used (12, 13).

Many internationally accepted reporting standards exist for different types of research designs, and provide empirically proven elements that need to be reported to improve the validity and transparency of research reports (14). Researchers need to be aware of these reporting standards and incorporate these elements into their research protocols, in order to improve the current state of research reporting. The evidence from articles published in leading journals in India, and from Indian trials reported in PubMed-indexed journals, suggests that this is often not the case (15, 16). This raises considerable uncertainty as to the reliability of their results.

5. Deception by design

Seemingly well designed, executed, and reported, RCTs with exciting results can also be misleading due to the hijacked research agenda. These trials are designed to deceive and the methods of deception are alarmingly simple, but effective. The main tactics used relate to the choice of comparators, the choice of outcomes, and the manipulation of statistics to produce desired outcomes, and selectively report them.

Lack of equipoise

One of the fundamental ethical principles underlying the conduct of RCTs is the “uncertainty principle” (or “clinical equipoise”), whereby there must be substantial uncertainty about the treatments being compared to help clinical decisions (17). A corollary of this is that the methods used should produce valid and generalisable results to help resolve this uncertainty for clinicians and policy makers.

The RCT is ideally placed to provide this with a new or competing intervention when the expected benefits outweigh risks. If all trials were conducted in accordance with the uncertainty principle, roughly half of them would favour the experimental intervention. However, empirical evidence indicates that compared to trials funded by non-profit agencies, more published industry-funded trials, though well designed to minimise the risk of bias, favour their intervention rather than the control intervention (18). This suggests a violation of the uncertainty principle in the design and reporting of these trials.

Explanatory versus pragmatic trials: efficacy versus effectiveness

Of the hundreds of thousands of clinical trials that have been conducted, less than 5- 10% are estimated to provide information considered crucial by clinicians and policy makers to guide clinical practice and health decisions (19).

Most RCTs funded by industry and academia are designed to demonstrate if a new drug works, for licensing and marketing purposes. In order to maximise the potential to demonstrate a “true” drug effect, homogenous patient populations; placebo controls; very tight control over experimental variables such as monitoring, drug doses, and compliance; outcomes addressing short term efficacy and safety; and methods to minimise bias required by regulatory agencies are used to demonstrate if, and how, the drug works under ideal conditions (19).

Placebo-controlled trials need fewer participants to demonstrate the superiority of an intervention. While there are scientific and regulatory reasons to ensure superiority of a new drug over placebo, these trials need to be followed by trials against standard treatments, and have a more pragmatic design. The designs used in these explanatory trials also exclude patients with more severe disease states, and other co-morbid medical conditions, unlike what clinicians see in clinical practice.

Practical or pragmatic clinical trials are designed to provide evidence for clinicians to treat patients seen in day-to day clinical practice, and evaluate their effectiveness under “real-world” conditions. These trials use few exclusion criteria and include people with co-morbid conditions, and all grades of severity. They compare active interventions that are standard practice, and in the flexible doses and levels of compliance seen in usual practice. They utilise outcomes that clinicians, patients, and their families consider important, such as satisfaction, adverse events, return to work, and quality of life (19). Recommendations exist on their design and reporting (20), but such trials are rare. They are usually funded by non-profit agencies, and are more likely to preserve equipoise in their design.

Choice of comparators: avoiding head-to-head comparisons

Industry sponsored trials rarely involve head-to head comparisons of active interventions, particularly those from other drug companies, thus limiting our ability to understand the relative merits of different interventions for the same condition. Even if they do occur they are more likely to report results and conclusions favouring the sponsor’s product compared to the comparator drug (2, 3, 7). Non-industry sponsored comparisons, on the other hand, often show little difference between one drug or the other, be they antidepressants, or antibiotics; many even show little difference in response rates between antidepressants, or antibiotics, compared to placebo (7).

Inappropriate comparisons

Even if active interventions are compared in industry-sponsored trials, the research agenda has devised ways in which the design of such trials is manipulated to ensure superiority of the sponsor’s drug. If one wants to prove better efficacy, then the comparator drug is a drug that is known to be less effective, or used in doses that are too low, or used in non-standard schedules or duration of treatment. If one wants to show greater safety, then the comparator is a drug with more adverse effects, or one that is used in toxic doses. Follow up, also, is typically too short to judge effectiveness over longer periods of time (2, 7).

Choice of outcome measures: ensuring statistical significance in advance

The choice of outcome measures used often ensures statistically significant results in advance, at the expense of clinically relevant or clinically important results. Outcomes likely to yield clinically meaningless results include the use of rating scales (depression, pain, etc.). These scales yield continuous measures usually summarised by means and standard deviations, rather than the dichotomous measures clinicians use such as: clinically improved versus not improved. These rating scales, however extensively validated, are hardly ever used in routine clinical practice. A difference of a few points on these scales results in *statistically significant* differences (low p values), that have little *clinical significance* to patients.

Other outcomes commonly used are *surrogate outcomes*; outcomes that are easy to assess but serve only as proxy indicators of what ought to be assessed, since the real outcome of interest may take a long time to develop. These are mostly continuous measures that require smaller sample sizes (blood sugar levels, blood pressure, lipid levels, CD4 counts, etc.). These measures easily achieve statistical significance but do not result in meaningful improvements (reduction in mortality, reduction in complications, improved quality of life) in patients' lives, when the interventions are used (often extensively) in clinical practice (7).

The use of *composite outcomes*, where many outcomes (primary, secondary and surrogate outcomes) are clubbed together (e.g.: mortality, non-fatal stroke, fatal stroke, blood pressure, creatinine values, rates of revascularisation) as a single primary outcome, can also mislead. Such trials also require smaller sample sizes, and increase the likelihood of statistically significant results. However, if the composite outcome includes those of little clinical importance (lowered blood pressure, or creatinine values), the likelihood of real benefit (reduction in mortality, or strokes, or hospitalisation) and the potential for harm (increase in non-fatal strokes or all-cause mortality) are masked (7).

Deceptive analysis and interpretation of data

The traditional approach to determining the significance of differences in outcomes between two interventions has been the use of p values from statistical tests of significance. The use of p values can be deceptive since a p value <0.05 only tells us that the observed differences were not due to chance with more than 95% certainty. If one intervention is 50% more effective than another intervention, the p values using traditional tests of significance will range from 0.29 (denoting that differences were not statistically significant) if only five people were compared in each arm, to <0.0001 (denoting that differences were statistically highly significant) if 100 people were compared in each arm. Any observed difference between two groups, *no matter how small*, can be made to be '*statistically significant*', at *any* level of significance, by taking a sufficiently large sample. P values do not tell us how effective the intervention is and whether this supposed effect is clinically important (*statistical vs. clinical significance*).

The use of estimates of relative effects of interventions, such as relative risks (RR) and odds ratios (OR) with their 95% confidence intervals, provides estimates of relative magnitudes of the differences and whether these exclude chance, as well as if these differences were nominal or likely to be clinically important.

However, even relative risks can be misleading since they ignore the baseline risk of developing the event without the intervention. The absolute risk reduction (ARR) is the difference in risk of the event in the intervention group and the control group, and is more informative since it provides an estimate of the magnitude of the risk reduction, as well the baseline risk (the risk without the intervention, or the risk in the control group). Systematic enquiry demonstrates that on average, people perceive risk reductions to be larger and are more persuaded to adopt a health intervention when its effect is presented as relative risks and relative risk reduction (a proportional reduction) rather than as absolute risk reduction; though this may be misleading (21).

Another statistical trick used to present favourable outcomes for interventions is the use of spurious *subgroup* analyses, where observed treatment effects are evaluated for differences across baseline characteristics (such as sex, or age, or in other subpopulations). While they are useful, if limited to a few biologically plausible subgroups, specified in advance, and reported as a hypothesis for confirmation in future trials; they are often used in industry-sponsored trials to present favourable outcomes, when the primary outcome(s) are not statistically significant.

Distorted or selective reporting

A considerable body of work provides direct empirical evidence that studies that report positive or significant results are more likely to be published; and outcomes that are statistically significant have higher odds of being fully reported, particularly in industry funded trials (4, 22). Evidence also shows that the published reports are not always consistent with their protocols, in terms of outcomes, as well as the analysis plan, and this again is determined by the significance of the results (4). Harms are very poorly reported in trials compared to results for efficacy; and are also often suppressed or minimised. Prospective trials registration was mooted as a deterrent against publication bias, and selective reporting, and has only partly succeeded, since not all trials are registered, and evidence of selective reporting continues even with pre-registered trials.

6. Let the buyer beware: conflicts of interest, ghost-writing and the marketing of research

Other tactics used to influence evidence-informed decision making include *ghost-writing*, where pharmaceutical companies hire public-relations firms who “ghost-write” articles, editorials, and commentaries under the names of eminent clinicians; a strategy that was detected by one survey in 75% of industry-sponsored trials, where the ghost author was not named at all, and in 91% when the ghost author was only mentioned in the acknowledgement section (22). Detecting such conflicts of interest is difficult, since they are rarely acknowledged due to the secrecy that shrouds the nexus between academia and industry in clinical trials.

Industry-sponsored trials often place various constraints on clinical investigators on publication of their results; these publication arrangements are common, allow sponsors control of how, when, and what is published; and are frequently not mentioned in published articles. These constraints go against current publication standards, and much needs to be done to educate and empower academic institutions in ensuring that industry-sponsored clinical trial arrangements fall within acceptable parameters.

Other measures used by pharmaceutical companies to market their products in the guise of educational improvement include visits from pharmaceutical sales representatives and distribution of gifts to residents and other doctors, advertisements in journals and prescribing software, sponsorship of meetings, sponsored speakers with hidden conflicts of interest, mailed information, and direct-to-patients advertising. Evidence from observational studies indicates that, with rare exceptions, exposure to pharmaceutical company information is associated with either no improved effect on physicians’ prescribing behaviour or with adverse effects in the form of reduced quality of prescribing, increased frequency of prescribing of sponsors’ products, or increased costs (23). Educating residents and clinicians and regulation of industry-academia relationships is needed in order to ensure that these influences are not detrimental to patients and healthcare.

An even more novel method by which medical treatments are marketed is the manufacturing of disease, or *disease mongering*, where, in pursuit of profits, pharmaceutical companies recruit academics and patient consumer groups, to widen the boundaries of illness (24). Thus normal people with non-specific symptoms are informed that they indicate treatable diseases; social problems and personal problems are turned into medical problems; biochemical and physiological parameters are lowered to indicate sub-threshold or actual diseases that need medical treatment; risk-factors for a disease are treated as an actual disease; and “life-style drugs” are advertised and marketed as drugs that all normal people ought to be using. With medical science so distorted and influenced by industry sponsorship, it is no wonder that there appears to be so little evidence one can trust.

7. Research misconduct

The current definition of research misconduct includes “fabrication, falsification, or plagiarism in proposing, performing, or reviewing research or in reporting research results” (25). Numerous high profile examples exist of scientific work retracted due to allegations of scientific misconduct and there is a suggestion from surveys of retracted papers that the incidence of misconduct is increasing. Papers retracted for fraud (data fabrication or data falsification) represent a deliberate effort to deceive, and surveys of these papers reveal that these papers were more likely to be published in journals with a high impact factor, and over 50% were authored by repeat offenders with other retracted papers. Falsified papers are indistinguishable from non-fraudulent ones and are slow in being retracted, particularly when senior researchers were implicated. Such publications put people at risk, even if retracted and in one study, retracted papers of clinical trials were cited over 5000 times, showing how they influence subsequent research. The extent to which people are put at risk in these trials is large and over 28,000 subjects were enrolled in 180 retracted primary studies, and over 400,000 subjects were enrolled in 851 secondary studies which cited a retracted paper (26). Most data in published papers are seldom independently verified and though expert statistical techniques exist that can identify scientific fraud in data-sets from clinical trials, these are seldom used.

The various acts of deception listed in this article are not considered as research misconduct by current definitions. However, misinformation arising from biased publications is much more common than deliberate fraud. They also deceive, and put people at risk since they are not retracted, and even if disputed, they continue to be cited. Many well-conceived RCTs are subverted by investigators whose actions violate research ethics. Surveys of authors of biased papers and researchers indicate that, in the majority of instances, authors were aware of their deception, though perceptions varied on the magnitude of this deception or of its consequences (27).

Fixing responsibility: whose trial is it anyway?

It would be unwise to lay responsibility for untrustworthy evidence at the door of EBM as an approach; but it is true that one needs to rebuild the evidence-base in the light of the many deceptions and misinformation in current research evidence. It is also easy to lay the blame entirely on pharmaceutical companies who are clearly central to the hijacking of the research agenda. However, it would be a gross disservice to say that all research sponsored by drug companies is false or inaccurate, just as it would be inaccurate to state that all academic research, or that funded by public agencies, is free from bias or misinformation.

The collective responsibility for improving the quality of research evidence lies with individual investigators, the institutions that research is conducted in, research and ethics committees who approve and supposedly monitor research approved by them, the funders and regulators of research, medical journal editors and peer-reviewers who publish research, and the scientific

community who use these results. Unless the scientific community awakens to the ways in which bias and misinformation have become accepted ways in which research is proposed, conducted, and reported, and collectively devises methods to halt the rot, the evidence that is supposed to inform patient care will continue to be tainted.

However, medical journalists, consumer groups, activists and the lay public also need to be educated on the various ways in which biased research and deceptive designs can be differentiated from valid and ethical research, and to work with the scientific community to reclaim the research agenda to advance the interests of science and healthcare.

Conclusions

Evidence-based medicine continues to be a valid approach to informing health decisions and the methods of EBM have contributed to identifying biases in many research designs and studies that now inform developments in the way research is interpreted and used. Future reports in this journal will focus on methods by which evidence that can be trusted can be identified and used to reliably inform health decisions.

Competing interests: The author is a contributor to the Cochrane Collaboration (www.cochrane.org) and director of one of the 14 independent Cochrane Centers (www.cochrane-sacn.org) worldwide. He has received research funding, travel support, and hospitality from organisations that support evidence-based healthcare. Funding support: The author is a salaried employee of the Christian Medical College, Vellore.

References

1. Straus SE, Richardson WS, Paul Glasziou, Haynes RB. *Evidence-based medicine: how to practice and teach EBM*. 3rd ed. Edinburgh: Churchill Livingstone; 2005.
2. Ioannidis JPA. Why most published research findings are false. *PLoS Med*. 2005; 2(8): e124.
3. Lathyris DN, Patsopoulos NA, Salanti G, Ioannidis JP. Industry sponsorship and selection of comparators in randomized clinical trials. *Eur J Clin Invest*. 2010 Feb; 40(2):172-82.
4. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, Decullier E, Easterbrook PJ, Von Elm E, Gamble C, Gherzi D, Ioannidis JP, Simes J, Williamson PR. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One*. 2008 Aug 28;3(8):e3081
5. Rochon PA, Mashari A, Cohen A, Misra A, Laxer D, Streiner DL, et al. Relation between randomized controlled trials published in leading general medical journals and the global burden of disease. *CMAJ*. 2004 May 25;170:1673-7.
6. Hill KP, Ross JS, Egilman DS, Krumholz HM. The ADVANTAGE seeding trial: a review of internal documents. *Ann Intern Med*. 2008 Aug 19;149:251-8.
7. Ioannidis JP. Perfect study, poor evidence: interpretation of biases preceding study design. *Semin Hematol*. 2008 Jul;45:160-6.
8. Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. *Ann Intern Med*. 2011 Jan 4;154(1):50-5.
9. Oxman AD, Lavis JN, Fretheim A, Lewin S. SUPPORT Tools for evidence-informed health policymaking (STP) 17: Dealing with insufficient research evidence. *Health Res Policy Syst*. 2009 Dec 16;7(Suppl 1):S17.
10. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA*. 1992 Jul 8;268(2):240-8.
11. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*. 1996 May 11;312(7040):1215-8.
12. Tharyan P, Adhikari SD. Randomized controlled clinical trials: critical issues. *J Anaesth Clin Pharmacol*. 2007;23(3):231-40.
13. Higgins JPT, Altman DG, Sterne JAC, editors. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated 2011 Mar). The Cochrane Collaboration [Internet] 2011. [cited 2011 Aug 29] Available from: <http://www.cochrane-handbook.org>.
14. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG; Consolidated Standards of Reporting Trials Group. CONSORT 2010 Explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*. 2010 Aug;63(8): e1-37.
15. Tharyan P, Premkumar TS, Mathew V, Barnabas JP, Manuelraj. Editorial policy and the reporting of randomized controlled trials: a survey of instructions for authors and assessment of trial reports in Indian medical journals (2004-05). *Natl Med J India*. 2008 Mar-Apr;21(2):62-8.
16. Zhang D, Freemantle N, Cheng KK. Are randomized trials conducted in China or India biased? A comparative empirical analysis. *J Clin Epidemiol*. 2011 Jan;64(1):90-5.
17. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987 Jul 16;317(3):141-5.
18. Djulbegovic B, Lacey M, Cantor A, Fields KK, Bennett CL, Adams JR, Kuderer NM, Lyman GH. The uncertainty principle and industry-sponsored research. *Lancet*. 2000 Aug 19;356(9230):635-8.
19. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D; CONSORT group; Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 2008 Nov 11;337:a2390.
20. Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, Costiniuk C, Blank D, Schünemann H. Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database Syst Rev*. 2011 Mar 16; (3):CD006776.
21. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003 May 31;326(7400):1167-70.
22. Gotzsche PC, Hrobjartsson A, Johansen HK, Haahr MT, Altman DG, Chan AW. Ghost authorship in industry-initiated randomised trials. *PLoS Med*. 2007 Jan;4(1):e19.
23. Spurling GK, Mansfield PR, Montgomery BD, Lexchin J, Doust J, Othman N, Vitry AI. Information from pharmaceutical companies and the quality, quantity, and cost of physicians' prescribing: a systematic review. *PLoS Med*. 2010 Oct 19;7(10):e1000352.
24. Moynihan R, Heath I, Henry D. Selling sickness: the pharmaceutical industry and disease mongering. *BMJ*. 2002 Apr 13;324(7342):886-91.
25. US Department of Health and Human Services. Public Health Services Polices on Research Misconduct; Final Rule. Part III. 42 CFR parts 50 and 93. Federal Register 2005;70(94):28370-400.
26. Steen RG. Retractions in the medical literature: how many patients are put at risk by flawed research? *J Med Ethics*. 2011 May 17 (Epub ahead of print).
27. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One*. 2009 May;4(5):e5738.