Asian Biomedicine Vol. 5 No. 2 April 2011; 175 - 178 **DOI: 10.5372/1905-7415.0502.022**

Editorial

'Lies, damned lies, and statistics': pharmaceutical, device, and academic enterprises 'in the dock'*

Robert E. Dedmon
Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA.

*British term meaning 'on trial', and referring to the courtroom area where the accused sits during criminal proceedings.

The widely-quoted 'Lies, Damned Lies, and Statistics' refers to the influence numerical data has in our everyday lives, as well as on policy makers of all disciplines. It has been erroneously attributed to Benjamin Disraeli, but probably originated from others [1]. While numerical data and meaningful information derived are critical in any research enterprise, they are also subject to error, both unintended and fraudulent.

Recently, two widely-circulated articles in prominent lay journals [1, 2] have called attention to these concerns, and it behooves all of us in the academic and clinical practice communities, along with the pharmaceutical and medical device industries to review the issues and strive to correct whatever deficiencies exist. This is a daunting task. However, it is an effort we owe our patients in providing them with safe and effective care. Public trust depends on integrity and objectivity of all the players, a goal remaining for the future! However, the immediate priorities are erroneous reporting related to funding pressures and the pursuit of statistical significance ($p < 0.05$), along with variance of perception among investigators analyzing identical data.

The two lay articles are based in large part on the work of John P.A. Ioannidis, MD PhD, currently a professor of epidemiology at Ioannina College of Medicine in Greece, with other appointments elsewhere (Harvard, Tufts, Stanford, etc). His discourse focuses on 'bad science', erroneous reporting in clinical trials, and the problems related to winning research funding for such trials [3, 4]. His theme, 'Why Most Research Findings are False' has sounded an alarm of immense proportions. The second article focuses on the psychology of perception and our inability to be objective, even with the 'gold standard' of RCT's.

The first lay article, by David H. Freedman in the November 2010 *Atlantic*, 'Lies, Damned Lies, and Medical Science', goes back to 2001 and summarizes questions initially raised regarding choice of patients for appendectomy (Albanian immigrants vs. patients with Greek names to increase 'scalpel time!). Then Freedman describes his interview experience with Ioannidis's team at Ioannina Medical School and their controversial agenda, Contradiction and publication bias in reported Randomized Clinical Trials (RCT's) of therapeutic interventions. He also raised the issue of whether drug firms were determining their studies in biased fashion to achieve more desirable results, i.e., the problem was asking the wrong question or asking it in a biased manner. He estimated that 80% of non-randomized studies turned out to be wrong, but more concerning that 25% of RCT's are as well! Ioannidis's first article, published *in PLoS Medicine* in 2005 [3] became the most-downloaded article in the publication's history. Freedman also refers to well-known examples such as hormone replacement therapy to prevent heart disease and the plethora of vitamin publications that have proved erroneous, both in Ioannidis's view and in more recent studies [5, 6].

The second lay article [2], by Jonah Lehrer in the December 10, 2013 issue of the *New Yorker*, 'The Truth Wears Off', begins with a discussion of declining effectiveness of second-generation antipsychotics, the issue of replicability, and that many medical concepts enshrined in textbooks have proven unconfirmable or outright wrong. He then reviews the work of Jonathan Schooler, a professor at the university of California-Santa Barbara and other researchers on 'verbal

Correspondence to: Robert E. Dedmon, MD MPH FACP FACOEM, 333 Park Drive, Neenah Wisconsin 54956, USA. Email: drbobred@aol.com

overshadowing', where '*subjects shown a face and asked to describe it were much less likely to recognize the face when shown it later than those who had simply looked at it.*' The article then proceeds to discuss other psychological concepts and the difficulty in getting negative results published in journals (publication bias). He also refers to the statistical technique of constructing a funnel graph discriminating studies with larger sample size from those with smaller size to evaluated possible bias. Schooler then raises the problem of faulty study design- 'We're wasting too much time chasing after bad studies and underpowered experiments. Every researcher should have to spell out, in advance, how many subjects they're going to use, and what exactly they're testing, and what constitutes a sufficient level of proof…' While the declining effect of second-generation antipsychotics may be 'regression to the mean', the decline may be related to 'the decline of illusion', not only in medical research of other disciplines as well. The article also cites the Work of Ioannidis referenced above and the quest for funding/ statistical significance in clinical research.

However, the news is not all hopeless, just disturbing. In addition, it reminds us of the need for asking the right research questions, exercising better attention to study design, and reporting, and exercising restraint in statistical analysis. Declining effect of drugs should not come as a surprise, as evidenced by the ongoing challenge of drug resistance in treating falciparum malaria [7, 8] and the perplexing problem of antibiotic resistance [9].

In his PLoS paper [3], Ioannidis expresses the opinion that 'Claimed research findings may often be simply accurate measures of the prevailing bias' and makes the following suggestions for improving the situation:

1. Better-powered evidence, such as larger studies and lower-biased meta-analyses. Such evidence should be based on a higher degree of pre-test probability than is currently practiced.
2. When research is conducted by several teams, use all the data for analysis and avoid emphasizing the results from just one.
3. Instead of chasing statistical significance, consider the chances of testing a true relationship and the meaning of R-the pre-study odds- in designing experiments as shown in **Table 1**.

Table 1. PPV of research findings for various combinations of power (1 - β), ratio of true to not true relationships (R), and Bias (u)

1 – β	*R*	*u*	Practical Example	PPV
0.80	1:1	0.10	Adequately powered RCT with little bias and 1:1 pre-study odds	0.85
0.95	2:1	0.30	Confirmatory meta-analysis of good-quality RCTs	0.85
0.80	1:3	0.40	Meta-analysis of small inconclusive studies	0.41
0.20	1:5	0.20	Underpowered, but well-performed phase I/II RCT	0.23
0.20	1:5	0.80	Underpowered, poorly performed phase I/II RCT	0.17
0.80	1:10	0.30	Adequately powered exploratory epidemiological study	0.20
0.20	1:10	0.30	Underpowered exploratory epidemiological study	0.12
0.20	1:1,000	0.80	Discovery-oriented exploratory research with massive testing	0.0010
0.20	1:1,000	0.20	As in previous example, but with more limited bias (more standardized)	0.0015

The estimated PPVs (positive predictive values) are derived assuming $\alpha = 0.05$ for a single study.
RCT, randomized controlled trial.
DOI: 10.1371/journal.pmed.0020124.t004

Source: Ioannidis, JPA. PLoS Medicine, 2005 [3]. Published under the terms of the Creative Commons Attribution License for open-access journals.

In his 2005 JAMA paper [4], Ioannidis focused on highly-cited articles in widely-read medical journals, and further analyzed 49 highly reputed research findings of the past 13 years, based on what we now call 'impact factor'(also subject to financial conflict of interest [10]). Forty-five claimed effective interventions, and seven (16%) were contradicted in subsequent studies, seven others (16%) reported exaggerated affects, 20 (44%) were replicated, and 11 (24%) remained unchallenged. Two of the best-known contradicted studies are the Nurses' Health Study (subsequent studies show that estrogen increased CAD risk in post-menopausal women), and the Health Professionals Study (subsequent studies showed that vitamin E supplementation does not reduce CAD in men). He further concludes that 'even the most highly cited randomized trials may be challenged and refuted over time, especially small ones'.

A few additional suggestions and issues have surfaced since these original 2005 reports, such as addressing missing data in clinical trials [11, 12] and the need to have full access to trial protocols and all results [13]. Other authors have modeled the impact of sequential analysis on decision-making in research [14] and the use of the Delphi technique to determine which outcomes to measure in clinical trials [15].

The question of objectivity was also recently emphasized in an article titled 'The interpretation of systematic reviews with meta-analysis: an objective or subjective process?' [16] The question they posed was on the efficacy of intravenous magnesium in the early post-myocardial infarction period. They organized the articles chronologically and separated them into packages with varying numbers of RCT reports and meta-analyses. They presented the packages to eight different reviewers, all of whom had published systematic reviews and/or meta-analyses, to answer three clinical questions: 1) Whether they believed magnesium is now proven beneficial; 2) Whether they believed magnesium will eventually be proven to be beneficial; and 3) Whether they would recommend its use at this time. Considerable disagreement ensued as the data became more heterogeneous for each package and each question. Some reviewers became more skeptical of its value and some became less skeptical over time. The authors concluded: 'the interpretation of the results of systematic reviews and meta-analyses includes a subjective component that can lead to discordant conclusions that are independent of the methodology used to obtain or analyze the data'.

What does all this tell us? I suggest:

1. Consider carefully the pre-test probability for a given outcome of an experiment.
2. Be sure to ask the right questions, and stay focused on these questions in conducting a trial.
3. For industry-sponsored drug trials, register at the outset, fully disclose protocols, and include all participants in analyses. Be careful about over-use of data mining software. Stop the fa□ade of dressing up your results [17].
4. Always reveal fully your funding sources and potential conflicts of interest.
5. Follow the CONSORT 2010 statement regarding clinical trials reporting [18].
6. Do a better job of graphical presentation of your results [19, 20].
7. For large trials across multiple countries and cultures, be mindful of the cultural, ethical, and scientific implications of such trials [21].
8. Publish negative results – editors, please note [22].

Finally, a brief comment about fraud in medical research: it must not be tolerated! The recent exposure of the Wakefield/autism debacle [23, 24] serves as a reminder of our responsibility as physicians, researchers, and teachers to be vigilant and set a good example for those who follow. Edmund Burke's oft-quoted statement serves as a blunt reminder: 'All that is necessary for the triumph of evil is for good men to do nothing'.

The author has no conflict of interest to declare.

References

1. Freedman DH. Lies, Damned Lies, and Medical Science. The Atlantic [Internet]. November 2010 [cited 2011 Jan 25]. Available from: http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/8269/
2. Lehrer J. The Truth Wears Off. Is there something wrong with the scientific method? The New Yorker [Internet]. 2010 Dec 13 [cited 2011 Jan 25]; Available from: http://www.newyorker.com/reporting/2010/12/

13/101213fa_fact_lehrer
3. Ioannidis JPA. Why most published research findings are false [Internet]. 2005 [cited 2011 Jan 25]; 2(8):124. Available from: http://www.plosmedicine.org
4. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. JAMA. 2005; 294(2):218-28.
5. Editorial. Anticancer vitamins du Jour-The ABCED's so far. Am J. Epidemiol. 2010; 172:1-3.
6. Stolzenberg-Solomon RZ, Jacobs EJ, Arsian AA, Qi D, Patel AV, Helzisouer KJ, et al. Circulating 25-hydroxyvitamin D and risk of pancreatic cancer. Am J. Epidemiol. 2010; 172:81-93.
7. Campbell CC. Malaria Control-Addressing challenges to ambitious goals. NEJM. 2009; 361(5):522-3.
8. World Health Organization. Global plan for artemisinin resistance containment [Internet]. 2011. [cited 2011 Jan 30]. Available from: http://www.who.int/malaria/publications/atozartemisinin_resistance_containment_2011.pdf)
9. Andersson DI, and Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? [Internet] 2011. [cited 2011 Jan 29]. Available from: http://www.nature.com/nrmicro/journal/v8/n4/abs/nrmicro2319.html
10. Lundh A, Barbateskovic M, Hrobjartsson A, Gotzsche PC. Conflicts of interest at medical journals; the influence of industry-supported randomized trials on journal impact factors and revenues – cohort study [Internet]. 2010 [cited 2011 Jan 26]; 7 (10): e1000354. Available from: http://www.plosmedicine.org
11. Fleming TR. Addressing missing data in clinical trials. Ann Intern Med. 2011; 154:113-7.
12. Smyth RMD, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. BMJ. 2010; 341:c7153.
13. Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results [Internet]. 2008 [cited 2011 Jan 25]; 5 (11): 230. Available from: http://www.plosmedicine.org
14. Pfeiffer T, Rand DG, Dreber A. Decision-making in research tasks with sequential testing [Internet]. 2009 [cited 2011 Jan 25]; 4 (2): 4607. Available from: http:// www.plosone.org
15. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies [Internet]. 2011 [cited 2011 Jan 26]. Available from: http://www.plosmedicine.org
16. Shrier I, Boivin JF, Platt RW, Steele RJ, Brophy JM, Carnevale F, et al. The interpretation of systematic reviews with meta-analyses: an objective or subjective process? [Internet]. BMC medical informatics and decision making. 2008 [cited 2011 Jan 28]; 8(19). Available from: http://www.biomedcentral.com/1472-6947-8-19
17. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation [Internet]. 2008 [cited 2011 Jan 28]; 5 (11):e217. Available from: http://www.plosmedicine.org
18. Wikipedia the Free Encyclopedia. [Internet]. Randomized Controlled Trial. [updated 2011 Jan 25; cited 2011 Jan 31]. Available from: http://en.wikipedia.org/wiki/Randomized_controlled_trial
19. Pocock SJ, Travison TG, Wruck LM. Figures in clinical trial reports: current practice and scope for improvement [Internet]. 2011. [cited 2011 Jan 26]. Available from: www.trialsjournal.com/content/8/1/36
20. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2011; 64:163-71.
21. Glickman SW, McHutchison JG, Petersen ED, Cairns CB, Harrington RA, Califf RM, Schulman KA. Ethical and scientific implications of the globalization of clinical research. NEJM. 2009; 360(8):816-23.
22. Journal of Negative Results in Biomedicine. [cited 2011 Jan 28]. Available from: http://www.jnrbm.com
23. Godlee F, Smith J, Marcovitch H. Wakefield's article linking MMR vaccine and autism was fraudulent. [cited 2011 Jan 15]. Available from: http://www.bmj.com/content/342/bmj.c7452.full.
24. Deer B. How the case against the MMR vaccine was fixed. [cited 2011 Jan 15]. Available from: http://www.bmj.com/content/342/bmj.c5347