

Washington University School of Medicine Digital Commons@Becker

Open Access Publications

2015

How do you know it is true? integrity in research and publications: AOA critical issue

Joseph A. Buckwalter

University of Iowa Hospitals and Clinics

Vernon T. Tolo

Keck School of Medicine at University of Southern California

Regis J. O'Keefe

Washington University School of Medicine in St. Louis

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation

Buckwalter, Joseph A.; Tolo, Vernon T.; and O'Keefe, Regis J., "How do you know it is true? integrity in research and publications: AOA critical issue." *The journal of bone and joint surgery*.97,1. e2. (2015).
http://digitalcommons.wustl.edu/open_access_pubs/3663

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.



THE AMERICAN ORTHOPAEDIC ASSOCIATION

Leadership in Orthopaedics since 1887

AOA Critical Issues

How Do You Know It Is True? Integrity in Research and Publications

AOA Critical Issues

Joseph A. Buckwalter, MS, MD, Vernon T. Tolo, MD, and Regis J. O'Keefe, MD, PhD

High-quality medical care is the result of clinical decisions based upon scientific principles garnered from basic, translational, and clinical research. Information regarding the natural history of diseases and their responses to various treatments is introduced into the medical literature through the approximately one million PubMed journal articles published each year. Pharmaceutical and device companies, universities, departments, and researchers all stand to gain from research publication. Basic and translational research is highly competitive. Success in obtaining research funding and career advancement requires scientific publication in the medical literature. Clinical research findings can lead to changes in the pattern of orthopaedic practice and have implications for the utilization of pharmaceuticals and orthopaedic devices. Research findings can be biased by ownership of patents and materials, funding sources, and consulting arrangements. The current high-stakes research environment has been characterized by an increase in plagiarism, falsification or manipulation of data, selected presentation of results, research bias, and inappropriate statistical analyses. It is the responsibility of the orthopaedic community to work collaboratively with industry, universities, departments, and medical researchers and educators to ensure the integrity of the content of the orthopaedic literature and to enable the incorporation of best practices in the care of orthopaedic patients.

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

The misrepresentation of natural observation has existed for as long as scientific research has been recorded^{1,2}. Ptolemy, the renowned second-century Egyptian astronomer, recorded

astronomical measurements that he could not have made. Ptolemy's work, purporting to prove that Earth was the center of the universe, influenced science and philosophy for centuries.

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. In addition, one or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

Copernicus, who revolutionized our understanding of both Earth and man's place in the universe, was accused of heresy when he reported a conflicting celestial configuration based on appropriate scientific methods and accurate measurements. The legendary physicist and Nobel laureate Robert Millikan (1868-1953), who discovered the negative charge of the electron, selected only fifty-eight of 140 observations for inclusion in his scientific presentations. While this selective use of data likely improved precision and the credibility of his claims, it did not truly represent his actual scientific findings. Sir Cyril Burt (1883-1971), a noted British psychologist, fabricated (extrapolated) data to show that human intelligence is 75% inherited. His work influenced educational programs and policies for generations.

Defining Scientific Misconduct: From the Obvious to the Subtle and Insidious

The U.S. Office of Research Integrity defines misconduct as “fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results” (<http://ori.hhs.gov/definition-misconduct>). Fabrication involves the presentation of observations or events that in fact never occurred; the experiments were never performed. Fabrication continues to occur and in some cases has influenced the treatment of patients with musculoskeletal disease. From 1996 to 2008, Dr. Scott Reuben published a series of articles that examined the potential role of cyclooxygenase-2 (COX-2) specific inhibitors in controlling postoperative pain following orthopaedic surgery. In a series of carefully designed and double-blind placebo-controlled studies, Dr. Reuben established that Celebrex (celecoxib; Pfizer), Bextra (valdecoxib; Pfizer), and Vioxx (rofecoxib; Merck) dramatically improved pain management for patients undergoing joint replacement, spine fusion, and anterior cruciate ligament reconstruction and decreased the complications associated with the standard use of opiates³. Dr. Reuben, a Professor of Anesthesiology and Pain Medicine at Tufts and the Chief of Acute Pain at Baystate Medical Center, was widely recognized for revolutionizing pain management for orthopaedic patients. A 2007 editorial in *Anesthesia & Analgesia* stated that Reuben had been at the “forefront of redesigning pain management protocols” through his “carefully planned” and “meticulously documented studies⁴.”

In 2008, it was discovered that two abstracts submitted by Dr. Reuben for Baystate Medical Center's Annual Research Week lacked institutional review board approval. Investigation showed that Dr. Reuben had never enrolled patients or performed the studies described in the manuscripts. Further review resulted in Baystate requesting medical journals to retract a combined total of twenty-one of Dr. Reuben's papers. Dr. Reuben's advocacy for COX-2 inhibitors to treat postoperative pain appeared in reviews, textbooks, and practice guidelines. Beginning in 2000, Reuben advocated that physicians should shift from the use of first-generation nonsteroidal anti-inflammatory drugs to the use of Vioxx, Celebrex, and Bextra to treat musculoskeletal pain³. Reuben urged the United States Food and Drug Administration (FDA) not to

restrict use of the drugs he studied, citing their efficacy and safety. Drug companies organized educational programs and symposia on the basis of Reuben's reports. Various editorials noted that “millions of orthopaedic patients' pain management has been affected by Dr. Reuben's research” and “Reuben's studies led to the sale of billions of dollars of Celebrex and Vioxx.”⁵

In 2010, Reuben, who had research supported by Pfizer and who served on a speaker's panel for the company, was sentenced to six months in prison, was ordered to pay a \$5000 fine, and was required to provide \$360,000 in restitution to drug companies. To many, it was surprising that Reuben could perform such extensive fabrication of high-impact research for so long. Reuben had numerous coauthors and worked in an academic institution where he successfully progressed through the promotion process⁵. These accounts demonstrate the high level of trust afforded to scientists and reveal that, although institutions approve research, they rarely monitor the findings associated with that research⁵.

Falsification involves the modification of scientific data so that it supports a particular hypothesis. In 1998, Dr. Andrew Wakefield and coauthors published a study in *The Lancet* of twelve children, suggesting a link between the measles, mumps, and rubella (MMR) vaccine and autism⁶. The results were widely reported by the media, were popularized on a variety of web sites, resulted in the refusal of vaccination by many parents, and led to lawsuits by parents of autistic children against vaccine manufacturers. *The Lancet* and the press later learned that Wakefield had received a \$110,000 payment from the Legal Aid Board prior to publishing the paper. The Legal Aid Board was seeking evidence that could be used in lawsuits against vaccine manufacturers and, following publication of the article, provided an additional \$674,000 payment to Wakefield. A retrospective review of the data used by Wakefield revealed that the diagnosis and/or dates of records were changed for all twelve children in the publication report so as to support the author's conclusions^{7,8}.

The Lancet partially retracted Wakefield's paper in 2004, and later issued a full retraction. The General Medical Council of the United Kingdom (U.K.) found Wakefield guilty of professional misconduct and revoked his medical license. However, public suspicion that vaccinations can cause autism persists. Vaccination rates have dropped sharply in many countries, including the United States, and this drop in vaccinations is a major contributor to the increased incidence of measles and mumps, resulting in outbreaks of the diseases and deaths in multiple countries⁹. Subsequent studies have demonstrated no link between the MMR vaccine and autism. Position statements supporting vaccination and the absence of a link with autism have been released by the Centers for Disease Control and Prevention (CDC), the American Academy of Pediatrics, the Institute of Medicine, the National Academy of Sciences, and the U.K. National Health Service. Nonetheless, the general public maintains a widespread belief in such an association. In 2012, the CDC reported a U.S. outbreak of whooping cough that infected 41,000 children—the largest outbreak since 1955.

This is an instance in which falsification has compromised public health worldwide.

Plagiarism is the inappropriate use of previously published information without attribution and with representation that the work is original. In 2011, an investigator reviewing the published literature (the National Library of Medicine's PubMed) related to osteoarthritis discovered an article that had been published twice in its entirety. The first instance of publication was in 2006, when the *Journal of Orthopaedic Research* published the paper with the title, "Chondrocyte Gene Expression in Osteoarthritis: Correlation with Disease Severity"¹⁰. In 2011, that same article was published as "Alterations in Expression of Cartilage-Specific Genes for Aggrecan and Collagen Type II in Osteoarthritis" in the *Romanian Journal of Morphology and Embryology*¹¹. The articles were identical except for the titles and list of authors. Dr. Mogoanta, Editor of the *Romanian Journal of Morphology and Embryology*, withdrew the article, notified PubMed, banned the authors from future publication, ceased collaboration with the reviewers, notified the Dean of Medicine and Chair of the Ethics Committee, and coauthored an editorial for the *Journal of Orthopaedic Research*¹².

While the case of plagiarism is clearly against the rules, stretching the boundaries of research ethics can be a more subtle and insidious process. More subtle events include an intentional failure to acknowledge previous work, intentionally incomplete or inaccurate description of methods, and repeat publication of similar work. In 2011, the *Journal of Orthopaedic Research* published an editorial that was titled "Publishing the Results of Multiple Experiments Using the Same Methods and Outcome Measures."¹³ The editorial described a scenario that occurred in 2002 and 2003 in which one author conducted a series of experiments using the same methods, and a single control group, to examine the effects of twenty-four different materials on bone formation. The author subsequently reported positive results in eleven publications, in ten different journals, over a period of seven years. In some manuscripts, the author failed to cite the previous publications, used the same description of the methods, and published similar or nearly identical figures. The same control group was used in each study, but this fact was not reported in the multiple publications.

Ensuring Accuracy in Research

The disclosure of conflicts of interest in publications and presentations has become standard. However, standards at different journals vary and typically do not involve detailed financial information. A study of disclosure at the 2012 meeting of the American Academy of Orthopaedic Surgeons revealed that 90% of the presentations had the required disclosure slide, but 10% did not¹⁴. The average amount of time dedicated to the disclosure slide was 3.1 seconds, permitting only a cursory review of the information. Only 45% of the disclosure slides had conflict-of-interest information that included coauthors. Only 15% of the disclosure slides had information regarding institutional conflicts¹⁴.

In a recent study in *The Journal of Bone & Joint Surgery*, a significant association was found between the funding source and the qualitative conclusions in publications examining the

prevention of deep vein thrombosis following total joint arthroplasty¹⁵. From 2004 to 2010, sixty-six studies published in PubMed-cited literature with identified industry or nonindustry funding were reviewed to determine whether a treatment was classified as favorable, neutral, or unfavorable in the prevention of clotting. Only two of the fifty-two industry-sponsored studies had negative results, in contrast to negative findings being observed in three of the fourteen studies without industry support¹⁵. A study of articles published in the MEDLINE database between 1980 and 2002 revealed a significant relationship between industry sponsorship and pro-industry conclusions¹⁶. The odds ratio was 3.6. Khan et al. reported that in five major journals, between 2002 and 2004, there existed a strong statistical link between industry funding and favorable outcomes¹⁷. While these reports do not necessarily indicate wrongdoing, they suggest possible increased risk of bias when interpreting results from industry-sponsored clinical trials.

Many journals now include information about levels of evidence based on trial design, and there has been a shift toward more publications being conducted with a higher level of evidence. The highest level of evidence is the placebo-controlled double-blind clinical trial with appropriate power analysis and statistical support. The lowest level of evidence involves the editorial or expert-opinion type of article (Fig. 1). The level of evidence supporting presentations at the American Academy of Orthopaedic Surgery meeting progressively increased from 2001 to 2010¹⁸. The number of studies utilizing Level-I or Level-II evidence has increased from 17% to 36% over that time (Fig. 2)¹⁸. In an article published in 2008, Okike et al. stated that their review of more than 1100 articles revealed that the major scientific factor influencing acceptance for publication in *The Journal of Bone & Joint Surgery* involved the level of evidence of the article¹⁹.

Journals now use information technology to help ensure the accurate and appropriate publication of research findings. Many instances of dual publication and/or plagiarism are identified by plagiarism-checking software. Similarities in language can be cross-referenced with all of the previously published materials and scientific journals. When a certain threshold of similarity is identified, a secondary review can be conducted.

Since 1975, there has been a tenfold increase in the number of journal articles that are retracted²⁰. The likelihood of retraction is greatest in high-impact journals that publish articles with substantial clinical relevance²⁰. A high-profile retraction of an article published in *The Journal of Bone & Joint Surgery* (British volume—now known as *The Bone & Joint Journal*) occurred in 2009^{21,22}. The manuscript showed that the use of bone morphogenetic protein-2 (BMP-2) in open tibial fractures in military service members improved healing and reduced the need for subsequent surgery^{21,22}. An inquiry arose after the journal received a report that the coauthors had not previously seen or approved the manuscript and that much of the paper was essentially false. Further investigation by the journal and by Walter Reed Army Medical Center (now known as the Walter Reed National Military Medical Center) led to retraction of the paper^{21,22}. Fraud is the leading cause of journal

TABLE American Academy of Orthopaedic Surgeons Levels of Evidence

| Levels of Evidence for Primary Research Question¹ | | | | |
|---|---|---|--|---|
| Types of Studies | | | | |
| | Therapeutic Studies— Investigating the Results of Treatment | Prognostic Studies— Investigating the Effect of a Patient Characteristic on the Outcome of Disease | Diagnostic Studies— Investigating a Diagnostic Test | Economic and Decision Analyses—Developing an Economic or Decision Model |
| Level I | <ul style="list-style-type: none"> High-quality randomized controlled trial with statistically significant difference or no statistically significant difference but narrow confidence intervals Systematic review² of Level-I randomized controlled trials (and study results were homogeneous³) | <ul style="list-style-type: none"> High-quality prospective study⁴ (all patients were enrolled at the same point in their disease with ≥80% follow-up of enrolled patients) Systematic review² of Level-I studies | <ul style="list-style-type: none"> Testing of previously developed diagnostic criteria in series of consecutive patients (with universally applied reference “gold” standard) Systematic review² of Level-I studies | <ul style="list-style-type: none"> Sensible costs and alternatives; values obtained from many studies; multiway sensitivity analyses Systematic review² of Level-I studies |
| Level II | <ul style="list-style-type: none"> Lesser-quality randomized controlled trial (e.g., <80% follow-up, no blinding, or improper randomization) Prospective⁴ comparative study⁵ Systematic review² of Level-II studies or Level-I studies with inconsistent results | <ul style="list-style-type: none"> Retrospective⁶ study Untreated controls from a randomized controlled trial Lesser-quality prospective study (e.g., patients enrolled at different points in their disease or <80% follow-up) Systematic review² of Level-II studies | <ul style="list-style-type: none"> Development of diagnostic criteria on basis of consecutive patients (with universally applied reference “gold” standard) Systematic review² of Level-II studies | <ul style="list-style-type: none"> Sensible costs and alternatives; values obtained from limited studies; multiway sensitivity analyses Systematic review² of Level-II studies |
| Level III | <ul style="list-style-type: none"> Case-control study⁷ Retrospective⁶ comparative study⁵ Systematic review² of Level-III studies | <ul style="list-style-type: none"> Case-control study⁷ | <ul style="list-style-type: none"> Study of nonconsecutive patients (without consistently applied reference “gold” standard) Systematic review² of Level-III studies | <ul style="list-style-type: none"> Analyses based on limited alternatives and costs; poor estimates Systematic review² of Level-III studies |
| Level IV | Case series ⁸ | Case series | <ul style="list-style-type: none"> Case-control study Poor reference standard | <ul style="list-style-type: none"> Analyses with no sensitivity analyses |
| Level V | Expert opinion | Expert opinion | Expert opinion | Expert opinion |

¹A complete assessment of the quality of individual studies requires critical appraisal of all aspects of the study design. ²A combination of results from two or more prior studies. ³Studies provided consistent results. ⁴Study was started before the first patient enrolled. ⁵Patients treated one way (e.g., with cemented hip arthroplasty) compared with patients treated another way (e.g., with cementless hip arthroplasty) at the same institution. ⁶Study was started after the first patient enrolled. ⁷Patients identified for the study based on their outcome (e.g., failed total hip arthroplasty), called “cases,” are compared with those who did not have the outcome (e.g., successful total hip arthroplasty), called “controls.” ⁸Patients treated one way with no comparison group of patients treated in another way. Reproduced from American Academy of Orthopaedic Surgeons. Levels of evidence for primary research question. <http://www.aaos.org/Research/Committee/Evidence/loetable1.pdf>. Accessed 2010 November 29.

Fig. 1
Levels of evidence in clinical trials. The highest level of evidence is the double-blind placebo controlled research design (Level I). The lowest level of evidence is the statement of expert opinion (Level V). (Table reproduced from Voleti PB, Donegan DJ, Baldwin KD, Lee GC. Level of evidence of presentations at American Academy of Orthopaedic Surgeons annual meetings. *J Bone Joint Surg Am.* 2012;94:e50[1-5].)

retractions; approximately 68% of the cases of retraction due to misconduct are the result of fraud, 43% are from duplicate publication, and 14% are due to plagiarism²⁰. Although the

number of retractions has markedly increased, these figures still likely tremendously underestimate fraudulent behavior. A confidential analysis of research behaviors published in *Nature*

| TABLE Percentage Difference in LOE by Presentation Type from 2001 to 2010 | | | | |
|---|-----------|------|------|-----------------------|
| Presentation Type | LOE | 2001 | 2010 | Percentage Difference |
| Poster | Level I | 1% | 5% | +4% |
| | Level II | 13% | 26% | +13% |
| | Level III | 18% | 31% | +13% |
| | Level IV | 68% | 38% | -30% |
| Paper | Level I | 3% | 9% | +6% |
| | Level II | 18% | 31% | +13% |
| | Level III | 28% | 35% | +7% |
| | Level IV | 51% | 25% | -26% |

Fig. 2
Differences in level of evidence in support of clinical research studies presented at the Annual Meeting of the American Academy of Orthopaedic Surgeons, 2001 to 2010. LOE = level of evidence. (Table reproduced from Voleti PB, Donegan DJ, Baldwin KD, Lee GC. Level of evidence of presentations at American Academy of Orthopaedic Surgeons annual meetings. *J Bone Joint Surg Am.* 2012;94:e50[1-5].)

in 2005 showed extensive questionable ethical behavior²³ (Fig. 3). Surprisingly, senior scientists were more likely to engage in questionable behaviors²³.

There were virtually no open-access journals in 1993. By 2009, more than 4500 open-access journals were in operation²⁴. The open-access journals are digital, online, free of charge, and in most cases lack copyright restriction. While advertised as “peer-reviewed” publications, the review process in many journals is cursory, and in some cases the review period is after publication. Publication in open-access journals is costly for authors. The Public Library of Science (PLOS) journals are leaders in the open-access publishing environment and have credibility. However, even the PLOS journals are very expensive, with publication cost to the author ranging from \$1300 to \$2900 per article, depending on the subspecialty journal. In 2011, the PLOS, BioMed Central, and Hindawi open-access journals collectively published over 50,000 articles²⁴. Thus, the open-access publishing business has become a multimillion-dollar business operation.

Despite safeguards, the likelihood that information entering the publication process will later be disproved is quite high. Dr. John Ioannidis has been a leader in evaluating the tendency for published materials to be later proven incorrect, incomplete, or inaccurate²⁵. He has published these findings both in scientific journals as well as in the popular press. His work suggests that essentially all Level-IV evidence reports will be proven false or inaccurate over time²⁵. Up to 40% of articles with Level-I evidence will also be disproved. He found that randomized clinical trials, when repeated, show similar results only two of every five times²⁵.

A New Paradigm: The Burden of Proof

There has been general advocacy for stronger actions and increased punishment for investigators participating in malfea-

sance in their research²⁶. The implications of research findings have amplified over time. The expenditure for drugs in 1990 was \$40 billion. In 2008, the cost of prescription drugs in the United States was \$234 billion, accounting for approximately 10% of the cost of U.S. health care²⁷. A recent publication has noted that \$3.6 billion was recovered in thirteen pharmaceutical fraud cases. Most of these cases were initiated by whistleblowers, and the pharmaceutical fraud cases collectively accounted for approximately 40% of all federal fraud whistleblower recoveries²⁷. As of July 2012, nine of the ten largest drug companies were bound by corporate integrity agreements under civil and criminal settlements or judgments in the United States²⁸. The *British Medical Journal* (now known as *The BMJ*) reported that companies systematically planned complex marketing campaigns to increase drug sales, including active promotion of off-label drugs or otherwise inappropriate use of drugs, despite knowledge that such use could seriously harm patients²⁸.

Positive research findings in the high-stakes pharmaceutical industry result in enormous financial rewards and offer new medical treatments. Successful drug research and medical treatments are innovative and at times push the boundary of public policy. An example is the potential for stem-cell therapy to result in improvements in tissue regeneration, to avert the process of aging, or to treat inflammatory diseases. Several highly notable cases of stem-cell fraud have gained public attention^{29,30}. Fraudulent reports can slow the progress of research, decrease the willingness of the public to fund scientific discovery, and greatly enhance general skepticism of the research community and its findings. Academic medical centers also bear increased risk in this era of scrutiny. After a recent case of fraud was discovered at Cornell Medical Center and Weill College of Medicine, Cornell University agreed to pay the government \$4.4 million to settle a Justice Department investigation³¹.

Technologies that are capable of examining large data sets have increased the amount of scientific scrutiny that is being performed. Search algorithms are available as a commercial contracting enterprise, in some cases are publicly available, and are sometimes proprietary research tools³². Both subspecialty journals and high-impact scientific journals now publish original articles that focus on large data sets that implicate individual investigators in research fraud³³⁻³⁵. Similarly, individual web sites make accusations and speculate about research fraud, and often publish this information on public web sites³⁶.

While “whistleblower” activities can enhance integrity in research, they are not foolproof. Investigators wrongly accused of research fraud or malfeasance face an extraordinarily adversarial environment. The burden of proof is not on the accuser but is on the investigator to validate their research findings and integrity. In many cases, the initial accusation is high profile with widespread coverage and interest. When the investigator is vindicated, the announcements are lower profile and frequently escape public notice. A claim of fraud for any investigator typically results in a formal review in an academic medical center, including the appointment of a committee to review the allegations. The investigator is required to retain and make available all of the original data and must report on the original data to the investigative

Table Percentage of scientists who say that they engaged in the behavior listed within the previous three years (n = 3,247)

| Top ten behaviours | All | Mid-career | Early-career |
|--|------|------------|--------------|
| 1. Falsifying or 'cooking' research data | 0.3 | 0.2 | 0.5 |
| 2. Ignoring major aspects of human-subject requirements | 0.3 | 0.3 | 0.4 |
| 3. Not properly disclosing involvement in firms whose products are based on one's own research | 0.3 | 0.4 | 0.3 |
| 4. Relationships with students, research subjects or clients that may be interpreted as questionable | 1.4 | 1.3 | 1.4 |
| 5. Using another's ideas without obtaining permission or giving due credit | 1.4 | 1.7 | 1.0 |
| 6. Unauthorized use of confidential information in connection with one's own research | 17 | 2.4 | 0.8*** |
| 7. Failing to present data that contradict one's own previous research | 6.0 | 6.5 | 5.3 |
| 8. Circumventing certain minor aspects of human-subject requirements | 7.6 | 9.0 | 6.0** |
| 9. Overlooking others' use of flawed data or questionable interpretation of data | 12.5 | 12.2 | 12.8 |
| 10. Changing the design, methodology or results of a study in response to pressure from a funding source | 15.5 | 20.6 | 9.5*** |
| Other behaviours | | | |
| 11. Publishing the same data or results in two or more publications | 4.7 | 5.9 | 3.4** |
| 12. Inappropriately assigning authorship credit | 10.0 | 12.3 | 7.4*** |
| 13. Withholding details of methodology or results in papers or proposals | 10.8 | 12.4 | 8.9** |
| 14. Using inadequate or inappropriate research designs | 13.5 | 14.6 | 12.2 |
| 15. Dropping observations or data points from analyses based on a gut feeling that they were inaccurate | 15.3 | 14.3 | 16.5 |
| 16. Inadequate record keeping related to research projects | 27.5 | 27.7 | 27.3 |
| Note: significance of χ^2 tests of differences between mid- and early-career scientists are noted by ** ($p < 0.01$) and *** ($p < 0.001$). | | | |

Fig. 3

Percentage of scientists admitting engaging in various questionable ethical behaviors at various career stages. (Reprinted by permission from Macmillan Publishers Ltd: [Nature]; table reproduced from Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature*. 2005 Jun 9;435[7043]:737-8.)

committee. Offices of public relations and legal departments also review these cases. A straightforward review process, given the time required, the number of senior officials involved in the process, and the multiple reports generated, typically costs between \$20,000 and \$50,000.

The research climate is high risk, particularly for investigators involved in industry-supported research involving drugs with the potential for widespread clinical use. An example is the approval of BMP-2 for use in spine fusions³⁷. During the FDA approval process, Medtronic supported seventeen clinical trials. These studies resulted in multiple publications and supported the use of BMP in spine fusion. However, additional evidence, accumulated over time, identified potential risk factors associated with BMP use in spine surgery. The original publications received ongoing analysis and scrutiny in the scientific literature and in the public press. Congressional investigation ensued and the orthopaedic surgeons involved in the studies were targeted in the press, in congressional hearings, and in the scientific literature³⁷. In response to these pressures,

Medtronic contracted with Yale University (Yale Open Data Access project; YODA)^{38,39}. Yale was given complete access to all of the original data obtained in each of the Medtronic-sponsored clinical trials. It is believed to be the first time individual patient data were made available in aggregate^{38,39}. To eliminate conflict of interest, Yale University served as a repository for the data but subcontracted two other universities, the Oregon Health & Sciences Center and the University of York in Great Britain, to perform an analysis of the individual patient data^{40,41}.

These reports were published in June of 2013^{40,41}. Both groups independently performed an analysis that included an aggregate of all individual patient data—an analysis that typically is not performed during the FDA approval process, but that provides increased sensitivity for the identification of rare complications. The analysis of the aggregate individual patient data was compared with a meta-analysis of the published clinical trials, and to a meta-analysis of the confidential clinical trials reports from each of the seventeen Medtronic-sponsored clinical trials that were submitted to the FDA. The aggregate

TABLE I Activities to Develop a Culture of Research Integrity

| | |
|-------------|---|
| Department | Establish research committees to approve studies and track progress Yearly department clinical research retreats Conferences, quality assurance, and grand rounds dedicated to research ethics Special review and approval process for industry-supported research |
| Institution | Educational programs in the ethics of research and publication Statistics cores to ensure accurate interpretation and presentation of data Masters and other degree-granting programs in clinical science designed for clinicians |
| Societies | Develop Continuing Medical Education (CME) courses regarding ethics in research and clinical care Symposia and conferences aimed at developing research expertise Develop strict criteria for the reporting of conflict of interest |
| Journals | Provide level of evidence for all published articles Reviewer training in publication ethics and statistical analysis Grading of statistics as a component of every article review Author and coauthor confirmation regarding the integrity of the data presented Development of an ethics board to provide final approval of manuscripts with conflict of interest |

data suggested a higher incidence of complications^{40,41}. The data further showed that while BMP is effective, it does not result in an incidence of spine fusion that is higher than what has been associated with traditional treatment, with use of autologous iliac crest bone graft^{40,41}. The YODA initiative suggests that the original conclusions and the published studies likely overestimated the benefit of BMP. Because this developed into a several billion dollar per year industry, the research was high-stakes, and skepticism developed regarding the integrity of the company and of the involved investigators, many of whom had financial relationships with Medtronic.

The Medtronic case has several important lessons. First, investigators conducting high-profile drug-related trials are at risk for intense scrutiny; second, companies involved in such studies may benefit from providing open access to the data and/or having the individual patient data from the individual trials aggregated; and third, the YODA initiative may provide a new model whereby the accuracy of clinical trials can be ensured, and individuals and academic institutions can be shielded from accusations of research fraud³⁹.

Conclusions

Why do authors engage in misconduct? Science, being a high-stakes enterprise, is based on the ability to produce new and important observations. An academic and/or industry scientific career is dependent on publication, which in turn has an impact on continued employment, promotion, grant support, personal recognition, and competition with other investigators. The current culture recognizes and celebrates discovery. However, the tortuous path to discovery with use of rigorous scientific principles, and the persistence required, are less appreciated. Good scientists are confident, and the burden of repeat experiments and the re-analysis of data can seem time-consuming and unnecessary. Research progress is sped up through deletion of “bad” data that can be rationalized and justified inappropriately. The victims are the scientific and clin-

ical community and the patients whom they serve. Research integrity remains a critical issue for the medical profession and for orthopaedics, and it is not simply a problem for the research community (Table I). The trends showing increasing instances of malfeasance in research can be reversed only with the sustained and collaborative effort of departments, medical centers or institutions, professional societies, and journals. It is essential that we develop a strong culture of ethical awareness—one that celebrates the integrity of the medical profession and the knowledge that guides patient care.

Source of Funding

This work was funded in part through a Public Health Services Award (P50 AR 054041) from the National Institutes of Health (R.J.O.). ■

Joseph A. Buckwalter, MS, MD
Department of Orthopaedics and Rehabilitation,
College of Medicine,
University of Iowa Hospitals and Clinics,
200 Hawkins Drive, 01008 JPP,
Iowa City, IA 52242

Vernon T. Tolo, MD
Children's Hospital Los Angeles,
Department of Orthopaedics,
Keck School of Medicine at University of Southern California,
1520 San Pablo Street, Suite 2000,
Los Angeles, CA 90033

Regis J. O'Keefe, MD, PhD
Department of Orthopaedic Surgery,
Washington University School of Medicine,
660 South Euclid, CB 8233,
St. Louis, MO 63110.
E-mail address: okeefe@wudosis.wustl.edu

References

1. Broad W, Wade N. *Betrayers of the truth: fraud and deceit in the halls of science*. New York: Simon and Schuster; 1983.
2. Goldstein D. *On fact and fraud: cautionary tales from the front lines of science*. Princeton and Oxford: Princeton University Press; 2010. p 168.
3. Reuben SS, Buvanendran A. Preventing the development of chronic pain after orthopaedic surgery with preventive multimodal analgesic techniques. *J Bone Joint Surg Am*. 2007 Jun;89(6):1343-58. Paper officially retracted in: Heckman JD. Retractions. *J Bone Joint Surg Am*. 2009 Apr;91(4): 965.
4. Goodman SB. Multimodal analgesia for orthopedic procedures. *Anesth Analg*. 2007 Jul;105(1):19-20.
5. White PF, Rosow CE, Shafer SL; Editorial Board of Anesthesia & Analgesia. The Scott Reuben saga: one last retraction. *Anesth Analg*. 2011 Mar;112(3):512-5.
6. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillo AP, Thomson MA, Harvey P, Valentine A, Davies SE, Walker-Smith JA. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*. 1998 Feb 28;351(9103):637-41. Paper partially retracted in: Editors of The Lancet. Retraction of an interpretation. *The Lancet*. 2004 Mar 6;363(9411):750. Paper fully retracted in: Editors of The Lancet. Retraction—Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*. 2010 Feb 6;375(9713):445.
7. Deer B. How the case against the MMR vaccine was fixed. *BMJ*. 2011;342:c5347. Epub 2011 Jan 5.
8. Deer B. Secrets of the MMR scare. How the vaccine crisis was meant to make money. *BMJ*. 2011;342:c5258. Epub 2011 Jan 11.
9. Cherry JD. Epidemic pertussis in 2012—the resurgence of a vaccine-preventable disease. *N Engl J Med*. 2012 Aug 30;367(9):785-7. Epub 2012 Aug 15.
10. Eid K, Thornhill TS, Glowacki J. Chondrocyte gene expression in osteoarthritis: correlation with disease severity. *J Orthop Res*. 2006 May;24(5):1062-8.
11. Jalbā BA, Jalbā CS, Vlādoi AD, Gherghina F, Stefan E, Cruce M. Alterations in expression of cartilage-specific genes for aggrecan and collagen type II in osteoarthritis. *Rom J Morphol Embryol*. 2011;52(2):587-91.
12. Buckwalter JA, Wright T, Mogoanta L, Alman B. Plagiarism: an assault on the integrity of scientific research. *J Orthop Res*. 2012 Dec;30(12):1867-8. Epub 2012 Aug 21.
13. Buckwalter JA, Wright TM, Donahue HJ, Amadio PC. Publishing the results of multiple experiments using the same methods and outcome measures. *J Orthop Res*. 2011;29:155-6.
14. Sassoon AA, Trousdale RT. Podium disclosures at the 2012 AAOS meeting: an exercise in going through the motions. *J Bone Joint Surg Am*. 2013 Apr 17;95(8):e51.
15. Lee YK, Chung CY, Koo KH, Lee KM, Ji HM, Park MS. Conflict of interest in the assessment of thromboprophylaxis after total joint arthroplasty: a systematic review. *J Bone Joint Surg Am*. 2012 Jan 4;94(1):27-33.
16. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA*. 2003 Jan 22-29;289(4): 454-65.
17. Khan SN, Mermer MJ, Myers E, Sandhu HS. The roles of funding source, clinical trial outcome, and quality of reporting in orthopedic surgery literature. *Am J Orthop (Belle Mead NJ)*. 2008 Dec;37(12):E205-12; discussion E212.
18. Voleti PB, Donegan DJ, Baldwin KD, Lee GC. Level of evidence of presentations at American Academy of Orthopaedic Surgeons annual meetings. *J Bone Joint Surg Am*. 2012 Apr 18;94(8):e50.
19. Okike K, Kocher MS, Mehlman CT, Heckman JD, Bhandari M. Publication bias in orthopaedic research: an analysis of scientific factors associated with publication in the *Journal of Bone and Joint Surgery (American Volume)*. *J Bone Joint Surg Am*. 2008 Mar;90(3):595-601.
20. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci U S A*. 2012 Oct 16;109(42): 17028-33. Epub 2012 Oct 1.
21. Kuklo TR, Groth AT, Anderson RC, Frisch HM, Islinger RB. Recombinant human bone morphogenetic protein-2 for grade III open segmental tibial fractures from combat injuries in Iraq. *J Bone Joint Surg Br*. 2008 Aug;90(8):1068-72. Paper retracted in Scott J. *J Bone Joint Surg Br*. 2009 Mar;91(3):285-6.
22. Scott J. Withdrawal of a paper. *J Bone Joint Surg Br*. 2009 Mar;91(3):285-6.
23. Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature*. 2005 Jun 9;435(7043):737-8.
24. Laakso M, Welling P, Bukvova H, Nyman L, Björk BC, Hedlund T. The development of open access journal publishing from 1993 to 2009. *PLoS One*. 2011;6(6): e20961. Epub 2011 Jun 13.
25. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005 Aug;2(8):e124. Epub 2005 Aug 30.
26. Face up to fraud. *Nature*. 2012 Jan 19;481(7381):237-8. Epub 2012 Jan 18.
27. Qureshi ZP, Sartor O, Xirasagar S, Liu Y, Bennett CL. Pharmaceutical fraud and abuse in the United States, 1996-2010. *Arch Intern Med*. 2011 Sep 12;171(16): 1503-6.
28. Davis C, Abraham J. Is there a cure for corporate crime in the drug industry? *BMJ*. 2013;346:f755. Epub 2013 Feb 6.
29. Cyranoski D. Stem-cell fraud hits febrile field. *Nature*. 2012 Oct 18;490(7420):321.
30. Snyder EY, Loring JF. Beyond fraud—stem-cell research continues. *N Engl J Med*. 2006 Jan 26;354(4):321-4.
31. Ready T. Cornell University scientists face charges of fraud. *Nat Med*. 2005 Aug;11(8):810.
32. Butler D. Journals step up plagiarism policing. *Nature*. 2010 Jul 8;466(7303):167.
33. Garner HR, McIver LJ, Waitzkin MB. Research funding: Same work, twice the money? *Nature*. 2013 Jan 31;493(7434):599-601.
34. Errami M, Sun Z, George AC, Long TC, Skinner MA, Wren JD, Garner HR. Identifying duplicate content using statistically improbable phrases. *Bioinformatics*. 2010 Jun 1;26(11):1453-7. Epub 2010 May 13.
35. Garner HR. Combating unethical publications with plagiarism detection services. *Urol Oncol*. 2011 Jan-Feb;29(1):95-9.
36. Couzin-Frankel J. Image manipulation. Author of popular blog that charged fraud unmasked. *Science*. 2013 Jan 11;339(6116):132.
37. Carragee EJ, Baker RM, Benzel EC, Bigos SJ, Cheng I, Corbin TP, Deyo RA, Hurwitz EL, Jarvik JG, Kang JD, Lurie JD, Mroz TE, Oner FC, Peul WC, Rainville J, Ratliff JK, Rihn JA, Rothman DJ, Schoene ML, Spengler DM, Weiner BK. A biologic without guidelines: the YODA project and the future of bone morphogenetic protein-2 research. *Spine J*. 2012 Oct;12(10):877-80.
38. Krumholz HM, Ross JS. A model for dissemination and independent analysis of industry data. *JAMA*. 2011 Oct 12;306(14):1593-4.
39. Krumholz HM, Ross JS, Gross CP, Emanuel EJ, Hodshon B, Ritchie JD, Low JB, Lehman R. A historic moment for open science: the Yale University Open Data Access project and Medtronic. *Ann Intern Med*. 2013 Jun 18;158(12):910-1.
40. Simmonds MC, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Rodgers MA, Stewart LA. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med*. 2013 Jun 18;158(12):877-89.
41. Fu R, Selph S, McDonagh M, Peterson K, Tiwari A, Chou R, Helfand M. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med*. 2013 Jun 18; 158(12):890-902.