

Detailed Case Histories

The following five detailed case histories of specific cases of actual and alleged research misconduct are included in an appendix to raise key issues and impart lessons that underlie the committee's findings and recommendations without breaking up the flow of the report. In several cases, including the translational omics case at Duke University and the Goodwin case at the University of Wisconsin, the committee heard directly from some of those involved.

The case histories differ in length in order to devote sufficient explanation to the issues involved in each case. For example, the translational omics case at Duke University unfolded over several years and involved multiple complex issues, making a lengthier discussion necessary. Issues covered in the cases include individual and institutional conflicts of interest, data falsification and fabrication, whistleblower retaliation and protection, insufficient or abusive mentoring, ghostwriting, authorship roles, institutional and administrator responsibilities, journal responsibilities, implementation of the federal government's research misconduct policy, and the costs and impacts of research misconduct.

Some cases mentioned in the report are not included in the appendix because the shorter descriptions already sufficed to illustrate the issues being described.

[Go to:](#)

THE WAKEFIELD MMR-AUTISM CASE

Synopsis and Rationale for Inclusion:*An undisclosed conflict of interest between a principal investigator and the entity funding their research can have far-reaching effects beyond the scope of the research study. In the MMR-autism case, Andrew Wakefield had undisclosed monetary conflicts of interest and was found to have violated human subjects protection rules in research underlying an article published in the Lancet ([UK GMC, 2010](#); [Triggle, 2010](#)).¹ In the opinion of the British Medical Journal, Wakefield also falsified data ([Godlee et al., 2011](#)). A formal retraction did not occur for over a decade, allowing ample time for the purported findings to become an important support for the anti-vaccine movement. This case not only confronts the issue of conflicts of interest but also weaknesses in institutional research governance, coauthor responsibility, and journal responsibility.*

In 1998, Andrew Wakefield published a paper in *The Lancet* claiming that he had found a link between the measles, mumps, and rubella (MMR) 3-in-1 vaccine and regressive autism, as well as a bowel disorder, using a sample of 12 children. Within a year, an article with a sample of 498 children rebutted Wakefield's findings, followed by additional rebuttal articles for several years thereafter ([Taylor et al., 1999](#)). However, Wakefield's article resonated with anti-vaccine movements in several countries, especially in the United Kingdom and United States, prompting some parents to refrain from vaccinating their children for fear of a connection to autism, contributing to decreased vaccination rates in the United States and United Kingdom and compromising the near success of eradicating these diseases from Western countries.

Six years after the 1998 article was published, 10 of the 12 coauthors retracted the paper's interpretation that the results suggested a possible causal link between the MMR vaccine and autism ([Murch et al., 2004](#)). In 2010, based on the UK General Medical Council's (GMC) Fitness to Practice Panel findings, *The Lancet* retracted the full article ([Lancet Editors, 2010](#)). Both of these retractions were prompted by the investigation by a British journalist, Brian Deer, initially published in the *Sunday Times* in early 2004. Deer exposed that Wakefield had

undisclosed financial interest in the research results, reporting that Wakefield had negotiated a contract with a lawyer who hired him to provide evidence against the MMR vaccine to help support a lawsuit against the MMR manufacturing company ([Deer, 2011a](#)). Deer reported that Wakefield profited approximately \$750,000 USD from the partnership ([Deer, 2011a](#)). In addition, Deer stated that Wakefield applied for a patent on his own measles vaccine, from which he was positioned to personally profit ([Deer, 2011a](#)). In addition, Deer reported that throughout the study, “Wakefield had repeatedly changed, misreported and misrepresented diagnoses, histories and descriptions of the children, which made it appear that the syndrome had been discovered” ([Deer, 2011a](#)). Lastly, Deer reported that the study sample was selectively recruited and not consecutively chosen as Wakefield had reported ([Deer, 2011a](#); [Wakefield et al., 1998](#), retracted). Deer then broadcast his findings on a UK television program, excerpts of which were later broadcast in the United States during an NBC *Dateline* investigation on Wakefield.

In addition to Deer's findings, the GMC found that Wakefield had performed unnecessary invasive tests on children that were “against their best interests,” was not qualified to perform the tests, did not have the necessary ethics approval to conduct his study, and unethically gathered blood samples by paying children at his son's birthday party for samples ([Triggle, 2010](#); [UK GMC, 2010](#)). He was found guilty of more than 30 charges of serious professional misconduct and removed from the UK's medical register ([Triggle, 2010](#); [UK GMC, 2010](#)).

Also in 2004 and soon after Deer's investigation, *The Lancet* launched an investigation of the paper. Other than undisclosed parallel funding and ongoing litigation, the *Lancet* reported that their editors did not find evidence of intentional deception or data falsification and so did not retract the paper ([Eggertson, 2010](#)). The article remained in the publication until the GMC's findings and subsequent actions in 2010, at which point *The Lancet* editors agreed “several elements of the 1998 paper by Wakefield et al are incorrect, contrary to the findings of an earlier investigation” and fully retracted the paper ([Lancet Editors, 2010](#)). The journal's editor, Richard Horton, said that “he did not have the evidence to [retract the paper] before the end of the GMC investigation” ([Boseley, 2010](#)).

In 2011, Brian Deer produced additional investigative reporting in support of his allegation that Wakefield falsified data, which was published by the British Medical Journal ([Deer, 2011b](#)). Deer's work was endorsed by the editors of BMJ ([Godlee et al., 2011](#)).

Wakefield denies ever having committed research misconduct; in a press complaint, Wakefield insisted “he never claimed that the children had regressive autism, nor that they were previously normal . . . never misreported or changed any findings in the study, never patented a measles vaccine . . . and he never received huge payments from the lawyer” ([Deer, 2011b](#)). Furthermore, he claims to be a victim of conspiracy via a Centers for Disease Control (CDC) cover-up, alleging the “CDC has known for years about an association between the MMR vaccine and autism” ([Ziv, 2015](#)). Wakefield's recent basis of this claim is a 2014 article by Brian Hooker published in *Translational Neurodegeneration* in which Hooker reevaluates data collected by the CDC and suggests African American boys who received the MMR vaccine before 24 months and after 36 months of age showed higher risks for autism ([Hooker, 2014](#), retracted). However, the Hooker paper was later retracted because of conflicts of interest and questionable research methods ([Translational Neurodegeneration Editor and Publisher, 2014](#)).

Following the 2004 investigation, Wakefield moved to the United States, where he is not licensed, but continues to defend the MMR-autism connection. He attempted to sue Deer and the *BMJ* in 2010 for defamation, but the lawsuit was dismissed ([Lindell, 2014](#)). Wakefield works out of Austin, Texas, as an anti-vaccine activist, where

he has received support from parents of children with autism ([Deer, 2014](#)). He directed the documentary *Vaxxed: From Cover-Up to Catastrophe*, which was to have been shown at the 2016 Tribeca Film Festival, but was withdrawn ([Goodman, 2016](#)).

In March 2011, the University College London (UCL), which took over the Royal Free Hospital where Wakefield worked at the time, announced intentions to conduct an institutional investigation on Wakefield ([Reich, 2011](#)). However, over 1 year later, UCL had not completed the investigation and explained that “given the passage of time, the fact that the majority of the main figures involved no longer work for UCL, and the fact that UCL lacks any legal powers of compulsion,” an investigation would not be a worthwhile endeavor for the university ([UCL, 2012](#)). Instead, UCL published a paper, *MMR and the Development of a Research Governance Framework in UCL*, detailing revisions made to the university's research governance framework in response to the shortcomings raised by the Wakefield case.

Go to:

PAXIL CASE

Synopsis and Rationale for Inclusion: *The Paxil case illustrates issues related to biomedical ghostwriting and unacknowledged conflicts of interest. In this practice, the listed authors of an article reporting on a clinical study may consist solely of prominent academicians, yet unacknowledged industry-supported researchers may have undertaken key tasks associated with the research, including aspects of concept design, subject enrollment, monitoring, data collection and interpretation, and writing the article. In extreme cases, the listed authors may not be able to confirm the integrity of the data or reported results. There have also been several notable cases over the past several decades in which suppression of negative findings or data falsification have been alleged or confirmed in industry-supported studies. Biomedical ghostwriting has been condemned by numerous scientific organizations worldwide.*

Ghostwriting, “the practice whereby individuals make significant contributions to writing a manuscript but are not named as authors,” has been condemned as an “example of fraud” and “a disturbing violation of academic integrity standards, which form the basis of scientific reliability” ([Bosch and Ross, 2012](#); [Stern and Lemmens, 2011](#)). The practice is not currently equated with plagiarism and so is not within the Office of Research Integrity's (ORI) power to regulate. [Bosch and Ross \(2012\)](#) suggest that ORI include ghostwriting in its definition of research misconduct so that it can be investigated and offenders can be punished under the federal research misconduct policy.

[ICMJE \(2015\)](#) established criteria against which to determine appropriate assignment of biomedical authorship and recommends that those who do not meet all of the criteria only be listed in the acknowledgments sections. [COPE \(2011\)](#) also recommends that specific rules be implemented to prevent ghostwriting, which is explicitly defined as misconduct in their guidelines.

If data are falsified or the reported results are misleading in a clinical study and the listed authors are not able to vouch for the integrity of the data or results, using the study as a basis for treating patients may present serious health and safety risks. If fabricated or falsified results are alleged for privately funded research, institutions are not required to report the investigation results to federal agencies under the federal research misconduct policy.

One example that illustrates these two issues is a 2001 paper overstating the benefits and understating the risks of the Glaxo SmithKline (GSK) drug Paxil in off-label treatment of children ([Basken, 2012](#)). Four GSK employees acted as whistleblowers, revealing “improper practices” to the U.S. government, including GSK

enticing doctors with vacations and knowingly publishing misreported data ([Thomas and Schmidt, 2012](#)). Although the lead authors listed on the paper were respected academics in the field, as part of Glaxo's \$3 billion settlement with the federal government, the company admitted that it had hired authors who were not listed as such and that the resulting publication had misrepresented the results.

Brown University, employer of the lead author, Martin B. Keller, launched an internal investigation, the results of which were not made public ([Basken, 2012](#)). No actions were taken against Keller, or the other 21 authors listed on the paper. Keller and at least five of the other authors continue to receive federal funding from the National Institutes of Health. *The Journal of the American Academy of Child and Adolescent Psychiatry*, which published the article, has not yet retracted it.

A recent reanalysis of Keller et al.'s 2001 study found no significant differences in efficacy between Paxil and the placebo in treating adolescents with major depression, but did find adverse emotional effects leading to increased suicidal thoughts and attempts for adolescents being treated with Paxil ([Le Noury et al., 2015](#)).

In 2015, Keller and 8 of the 22 authors of the original study wrote a letter to the blog *Retraction Watch* rebutting many points of [Le Noury et al.'s 2015](#) reanalysis of the study; Keller claimed that data used in the reanalysis were not available during the time of the original study. He also firmly asserted that none of the paper was ghostwritten. Keller concluded that describing the original “trial as ‘misreported’ is pejorative and wrong,” specifically from a retrospective point of view ([Keller et al., 2015](#)).

At this point, it appears that key issues related to this episode may never be resolved. In addition to the Paxil case, there have been several other cases of possible biomedical ghost writing that led to legal consequences for both medical companies and ghostwriters, indicating a heightened level of responsibility on the part of authors (see [Chapter 7](#)).

The Food and Drug Administration recently released draft guidance on publications reporting use of approved products for off-label indications: *Guidance for Industry Distributing Scientific and Medical Publications on Risk Information for Approved Prescription Drugs and Biological Products—Recommended Practices*. The guidelines state that scientific journals should not publish articles “written, edited, excerpted, or published specifically for, or at the request of, a drug or device manufacturer,” nor “be edited or significantly influenced by a drug or device manufacturer or any individuals having a financial relationship with the manufacturer” ([FDA, 2014](#)). In addition, articles including information on pharmaceuticals should include a statement disclosing the manufacturer's interest in the drug and any financial interest between authors and the manufacturer ([FDA, 2014](#)). Final guidance is expected, but has not yet been released.

Go to:

THE GOODWIN CASE AT THE UNIVERSITY OF WISCONSIN

Synopsis and Rationale for Inclusion: Graduate students may need support and protection from repercussions that may arise as a result of research misconduct committed by their mentor. Students stand to lose years of work if their mentor is found guilty of research misconduct, and may need to find another research group to continue their work, restart their graduate research from the beginning, or leave academia completely. With this in mind, graduate students of Elizabeth Goodwin, formerly a geneticist at the University of Wisconsin, found that data had been fabricated in one of Goodwin's proposals and reported her to the university. This case demonstrates difficult choices that may confront whistleblowers, especially those in vulnerable positions such as graduate students or postdoctoral fellows, the need for institutions to support young researchers put into

difficult situations through no fault of their own, and the need for better mentoring in some laboratory and institutional environments.

In fall 2005, graduate students working in the laboratory of University of Wisconsin geneticist Elizabeth Goodwin were confronted with evidence that their advisor had falsified data contained in a proposal to the National Institutes of Health ([Couzin, 2006](#)). Specifically, one experiment described in the proposal had not actually been performed, and figures appeared to have been manipulated. Over a period of several months, the students sought explanations from Goodwin, with which they were ultimately unsatisfied, and discussed among themselves what they should do ([Allen, 2012](#)). Recognizing that a decision to bring their concerns to university administrators would essentially shut down Goodwin's lab and have a severe negative impact on their own graduate careers, they decided that any such decision would need to be made unanimously.

Ultimately, the students decided to turn Goodwin in, which led to a university investigation finding that data in several grant applications had been falsified, a ruling confirmed by the Office of Research Integrity ([ORI, 2010](#)). Goodwin also pled guilty to making false statements on government documents, and was sentenced to 2 years' probation, fined \$500, and was ordered to pay \$100,000 in restitution ([Winter, 2010](#)). Several papers that Goodwin had coauthored were also investigated, but falsification was not found.

As they anticipated, the graduate students did suffer negative impacts from the case ([Allen, 2012](#)). One was able to continue work in another lab, and one was able to start a new project in a different lab at Wisconsin. One left Wisconsin to enter the PhD program at another institution, essentially starting over after 4 years. The remaining three students decided to embark on careers outside of academic research.

The case highlights several key issues. The first is the importance of whistleblowers to the system of ensuring research integrity. Although failure to replicate results, statistical analysis, and other mechanisms may be increasingly important in uncovering research misconduct, postdoctoral fellows and graduate students are responsible for reporting a significant percentage (up to half) of cases involving nonclinical research that come to ORI ([Couzin, 2006](#)). And these whistleblowers often suffer negative consequences, primarily severe damage to their careers, even when the institution takes appropriate steps to protect them from retaliation.

In addition, former students report that in the years immediately preceding Goodwin's falsified applications, problems were apparent in the lab. Several students were not making progress on their research, with no publications to show for years of work, but were advised to continue on these “dead projects” ([Allen, 2012](#)). Goodwin had also reportedly been encouraging students to overinterpret data and conceal data that conflicted with desired results ([Couzin, 2006](#)). Such ineffective mentoring and promotion of detrimental research practices create a poor environment for research integrity.

[Go to:](#)

THE HWANG STEM CELL CASE AND THE UNIVERSITY OF PITTSBURGH: COAUTHOR RESPONSIBILITIES AND INSTITUTIONAL RESPONSES

***Synopsis and Rationale for Inclusion:** The Hwang case raises several important research integrity issues, including data fabrication and falsification, abuse of mentorship status, whistleblower retaliation, and endangering the health of trial participants. The University of Pittsburgh's role in this case highlights the need for institutional oversight and defined standards for authorship roles. A second, more recent case at the University of Pittsburgh further demonstrates the need for oversight and institutional focus on addressing all cases of research misconduct.*

One highly publicized case that raises several important research integrity issues is that of Hwang Woo-suk, whose purportedly groundbreaking stem cell research turned out to be based on fabricated experiments ([Holden, 2006](#)). In his first article published in *Science* (in 2004), Hwang claimed to have “generated embryonic stem cells from an adult human cell,” a process often referred to as therapeutic cloning, so that cells could be transplanted “without immune rejection to treat degenerative disorders” ([Wade, 2006](#); [Hwang et al., 2004](#), retracted). University of Pittsburgh stem cell researcher Gerald Schatten began corresponding with Dr. Hwang in late 2003, offering editorial input and support to Hwang's 2004 paper that had earlier been rejected by *Science*. Following the acceptance of the paper, Schatten and Hwang began discussing a follow-up paper in which Hwang claimed his laboratory team had “created human embryonic stem cells genetically matched to specific patients” ([Sang-Hun, 2009](#)). According to Schatten, he and Hwang drafted and edited the article together; Schatten was responsible for much of the writing and was a prominent public promoter of the findings ([University of Pittsburgh, 2006](#)). The article was published in *Science* in 2005 naming Schatten as a senior author, a role he later denied, claiming to have been no more than a coauthor.

In June 2005, immediately following the second article's published release and Hwang's announcement of a clinical trial, Young-Joon Ryu, a former researcher in Hwang's laboratory aware of the fabricated data, worried for the safety of trial participants. Ryu e-mailed Korean television network, Munhwa Broadcasting Corporation (MBC) recommending an investigation ([Cyranoski, 2014b](#)). Unfortunately, Ryu endured negative effects for his role as a whistleblower. Ryu's identity was leaked early on in the MBC investigation and he received negative backlash from Hwang's ardent supporters that led to Ryu's resignation from his position at a hospital and to a period of unemployment.

As the MBC investigation was under way, ethical concerns with Hwang's research methods were being raised. Sun Il Roh, a coauthor of the 2005 paper and fertility specialist at a hospital in Seoul, disclosed that 20 eggs he had provided to Hwang for the study had been paid for (a violation of human subjects protections), but that Hwang was unaware of this ([Cyranoski and Check, 2005a](#)). Amongst this and other signs that accepted ethical procedures were not being followed, including that a young, female graduate student in Hwang's laboratory had donated eggs to the experiment (another violation of human subjects standards), Schatten asked that his name be removed from the 2005 publication and ceased working with Hwang ([Cyranoski and Check, 2005b](#)). Four days after Roh came forward and after a year of denials, Hwang admitted that “his stem-cell research used eggs from paid donors and junior members of his team” ([Cyranoski and Check, 2005a](#)). Days later, Hwang revealed to *Science* that of the 11 photos used in the 2005 article, several were duplicates, “even though each was meant to show a different human cell colony” ([Wade, 2005](#)). Hwang claimed that this was a mistake and that it occurred only when *Science* requested higher-resolution photos, not in the original submission. Roh was interviewed in the MBC television broadcast on Hwang and revealed that “Hwang had told him ‘there are no cloned embryonic stem cells’” ([Cyranoski, 2005](#)).

After its formal investigation in 2005, a Seoul National University committee determined that both of Hwang's articles were based on fabricated data ([SNU, 2006](#)). Numerous accusations ensued with Hwang admitting to “ordering subordinates to fabricate data,” but also blaming a coauthor who “admitted to switching stem cells without Hwang's knowledge” ([Cyranoski, 2014c](#)). Preceding the SNU investigation's conclusion, Schatten and Hwang had together requested that the paper be retracted from *Science*. Based on the investigation findings, Donald Kennedy, *Science* editor-in-chief, retracted both the 2004 and 2005 papers, reporting that “seven of the 15 authors of [Hwang et al., 2004](#) have agreed to retract their paper” and “all of the authors of Hwang et al., 2005 have agreed to retract their paper” ([Kennedy, 2006](#)). Following the retractions, Korea's National Bioethics Committee (created in response to ethical questions concerning Hwang's early research) found that Hwang had

“forced junior members of his lab to donate eggs, and that he used more than 2,221 eggs in his research” ([Nature, 2005](#)). Hwang had only reported using approximately 400 eggs. Throughout the entire investigation, Hwang maintained that his laboratory did “create stem cells matched to individual patients,” but acknowledged that mistakes were made throughout the research process. His achievement of the first cloned dog, Snuppy, was never discredited ([Nature, 2005](#)).

Hwang was indicted on three charges, “embezzling KRW2.8 billion [(US\$2.4 million)], committing fraud by knowingly using fabricated data to apply for research funds, and violating a bioethics law that outlaws the purchase of eggs for research” ([Nature, 2005](#)). In 2009, Hwang was convicted on two of the three charges, violating the bioethics law and embezzling government funds. The fraud charge was dropped because the “companies involved gave the money knowing that they would not benefit from the donation” ([Cyranoski, 2014a](#)). Hwang was sentenced to a 2-year suspended prison sentence.

Today, with private funding, Hwang runs the Sooam Biotech Research Foundation that he opened in July 2006. The laboratory clones animals with the goals of “producing drugs, curing diabetes and Alzheimer's disease, providing transplantable organs, saving endangered species and relieving grief-stricken pet owners” ([Cyranoski, 2014a](#)). Since opening Sooam, Hwang has been published in peer-reviewed journals and has been successful in obtaining a Canadian patent on a cloned cell line (NT-1), which was found to be fraudulent in Hwang's 2004 *Science* article. While Hwang attempts to make a comeback, he has twice been denied approval for therapeutic cloning of human embryos by the Korean health ministry and, for now, continues to clone animals.

While a subsequent investigation by a University of Pittsburgh panel found that Gerald Schatten had not been involved with the fabrication, the incident raised questions about whether Schatten's contributions to the paper merited authorship in the first place. To what extent should coauthors, honorary or otherwise, be held responsible for the fabricated results of their collaborators? Schatten argued over the definition of the term *write*, as he did not generate the data on which the text was based, but the panel found this and disagreements over the definition of *senior author* to be dishonest attempts to relieve himself of responsibility ([University of Pittsburgh, 2006](#)). The panel found Schatten's authorship role to be reasonable given that he wrote each draft of the paper. Schatten was also named coauthor on Hwang's 2005 Snuppy paper; however, Schatten reported to the panel that his “major contribution to the paper” was to suggest using a professional photographer to present Snuppy ([University of Pittsburgh, 2006](#)). The panel did not doubt this claim, but found it “less clear that this contribution fully justified co-authorship” ([University of Pittsburgh, 2006](#)). At his own request, Schatten was not acknowledged in Hwang's 2004 paper. Among questions of the appropriateness of authorship, also ethically problematic was Schatten's acceptance of approximately \$40,000 in honoraria and research proposals to Hwang's laboratory valued at more than \$200,000 for a 4-month period with implications that the grant would be continued annually ([University of Pittsburgh, 2006](#)).

The University of Pittsburgh panel's report stated that Schatten “did not exercise a sufficiently critical perspective as a scientist,” but because he likely did not “intentionally falsify or fabricate experimental data, and there is no evidence that he was aware of the misconduct,” he was found guilty of “research misbehavior” rather than “research misconduct” ([University of Pittsburgh, 2006](#)). “Research misbehavior” was not used or defined in the University of Pittsburgh research misconduct policy in effect at the time. The panel did not recommend any specific disciplinary action against him. Chris Pascal, director of the Office of Research Integrity supported the decision, stating “universities have a right to add refinements to categories of malfeasance” ([Holden, 2006](#)). The term *research impropriety* is contained in the University of Pittsburgh research misconduct policy adopted in 2008 ([University of Pittsburgh, 2008](#)).

THE TRANSLATIONAL OMICS CASE AT DUKE

Synopsis and Rationale for Inclusion: *The case of Duke University researchers Joseph Nevins and Anil Potti, which stretched out over several years and attracted national media attention, illustrates shortcomings and deficiencies in current approaches to research integrity on the part of researchers, research institutions, government agencies and journals ([CBS News, 2012](#)). Potti's fabricated results endangered trial participants and may have contributed to public mistrust in scientific research. Institutionally, supervisors at the laboratory level and senior administrators did not respond effectively for several years despite multiple warning signs. This case also raises questions about the responsibility of a journal to respond appropriately if numerous inquiries are made on the same original article. Several parties' unresponsiveness to questions on Potti's work may have delayed the findings of research misconduct.*

Omics is the study of molecules in cells, such as DNA sequences (genomics) and proteins (proteomics). Translational omics research seeks to apply this new knowledge to the creation of diagnostic tests that better detect disease and determine individualized treatment. Translational omics involves several significant challenges. Research “generates complex high-dimensional data” and resulting diagnostics are characterized by “difficulty in defining the biological rationale . . . based on multiple individual biomarkers” ([IOM, 2012](#)). In addition, diagnostic tests differ from drugs and other medical technologies regarding regulatory oversight; tests may be reviewed by the Food and Drug Administration, or be validated in a CLIA-certified laboratory (Clinical Laboratory Improvement Act).

Beginning in 2006, a series of papers appearing in major journals such as *Nature Medicine* and the *New England Journal of Medicine* purported to show that the gene activity in a patient's tumor cells could be used to determine which chemotherapy drugs would be most effective for that patient. This capability would enable significant advances in cancer treatment. Since individual reactions to these drugs are heterogeneous, the drugs that are effective for one person may not be effective for another. The lead author of the papers was cancer researcher Anil Potti, who worked at Duke University in the lab of Joseph Nevins.

Soon after the first papers were published, Keith Baggerly, Kevin Coombes, and Jing Wang, bioinformaticians at the M. D. Anderson Cancer Center of the University of Texas, began working to replicate the results. They immediately encountered difficulties using the data made publicly available with the paper, and began communicating with Potti and Nevins. Data provided by the Duke team to Baggerly, Coombes, and Wang contained numerous anomalies and obvious errors, making it impossible to replicate or verify the results. A correspondence by the M. D. Anderson researchers submitted to *Nature Medicine* in 2007 raising these issues was quickly rebutted by Potti and Nevins ([Coombes et al., 2007](#); [Potti and Nevins, 2007](#)). However, when Baggerly, Coombes, and Wang examined additional information provided by the Duke team they found that there were still significant problems. For example, in some cases, sensitive and resistant labels for cell lines were reversed, which would lead to patients being treated with the least effective chemotherapy drug if the tests were used to direct treatment, rather than the most effective.

Over the next several years, in response to interest expressed by M. D. Anderson clinicians in utilizing the advances that continued to be reported by Potti and Nevins, Baggerly and Coombes worked with the data. In several cases where they discovered clearly incorrect results, they submitted correspondence to journals such as *Lancet Oncology*, *Journal of Clinical Oncology*, and *Nature Medicine*, but these were rejected without explanation ([Baggerly, 2010, 2012](#)).

In 2007, at the same time questions were being raised about the data underlying the Nevins-Potti research, Duke University and Duke University Medical Center investigators not associated with Nevins or Potti launched three clinical trials based on the results, and an additional trial was launched at Moffitt Cancer Center ([IOM, 2012](#)). Duke also applied for patents, and several companies were working to commercialize the research, including one in which Potti served as a director and secretary ([Reich, 2010b](#); [Tracer, 2010](#)). Learning about the trials in June 2009, Baggerly and Coombes prepared a critical analysis of the Duke work, which was published in the *Annals of Applied Statistics* after it had been rejected by a biomedical journal ([Baggerly and Coombes, 2009](#)).

In January 2015, the *Cancer Letter*, a specialist newsletter, reported that Bradford Perez, a third-year medical student who was working with Potti in the Nevins lab, became very concerned about the methodology and reliability of the research ([Goldberg, 2015](#)). He shared these concerns in a detailed memo with Potti, Nevins, and several Duke administrators in the spring of 2008 ([Goldberg, 2015](#)). In addition to providing specifics about a number of concerning factors, he asked that his name be removed from four papers based on the work he had contributed to, including a paper submitted to the *Journal of Clinical Oncology*, and left the Nevins-Potti laboratory ([Perez, 2008](#)). Rather than catalyzing any independent assessment of the serious concerns raised by Perez about the quality of the research, Duke administrators referred him back to Nevins with no apparent follow-up by any institutional official. Nevins and Potti committed to revalidate all of their work, but it appears that this did not happen. Perez left the Nevins lab knowing he would repeat a year of his medical education, in his words, “to gain a more meaningful research experience” ([Perez, 2008](#)).

As noted in a 2012 Institute of Medicine (IOM) report discussed further below, Duke “did not institute extra oversight or launch formal investigations of the three trials during the first 3 years after the original publications triggered widely known controversy about the scientific claims and after concerns started to develop about the possible premature early initiation of clinical trials” ([IOM, 2012](#)). Not only did Duke's administration fail to act decisively on Perez's suspicions, but an administrator who counseled Perez on the matter did not even inform the IOM committee that Perez had come forward years earlier ([Goldberg, 2015](#); [IOM, 2012](#)). In response to the 2015 revelations by the *Cancer Letter*, Duke Medicine officials did not answer specific questions, but did state that “there are many aspects of this situation that would have been handled differently had there been more complete information at the time decisions were made” ([Goldberg, 2015](#)).

National Cancer Institute (NCI) researcher Lisa McShane had also been unsuccessful in attempts to replicate the work ([Economist, 2013](#)). In the fall of 2009, NCI expressed concern about the clinical trials at Duke as well as the parallel trial at Moffitt. The trials were suspended, and Duke's Institutional Review Board formed an external review panel to evaluate the concerns. The Duke trials were restarted in early 2010 after the review panel concluded that the approaches used in the trials were “viable and likely to succeed” ([IOM, 2012](#)).

During the first half of 2010, NCI continued to raise questions about the research. Through a Freedom of Information Act request *submitted by the Cancer Letter*, it was revealed that the external review panel was not provided with several critical pieces of information, including a detailed description of the statistical methods used in the original research, and a new critique from Baggerly and Coombes based on analysis of updated data posted by Potti and Nevins ([Baggerly, 2010](#); [Duke University, 2009](#)). About that material, the 2012 IOM report notes that it “was never forwarded to the external statistical reviewers because of the university leadership's concerns that it might ‘bias’ the committee's review” ([IOM, 2012](#)).

Several developments in July 2010 brought matters to a head. It was reported that Potti's claim on his resume that he had been a Rhodes Scholar was exaggerated, and this was confirmed by the University of Oxford

([Goldberg, 2010](#); [Singer, 2010](#)). Also, several dozen prominent biostatisticians wrote to NCI director Harold Varmus to request that the clinical trials based on the Duke research be suspended until the science could be publicly clarified ([Barón et.al., 2010](#); [Singer, 2010](#)). Duke suspended the trials and suspended Anil Potti's employment in response. The trials were ultimately terminated and Potti left Duke. Starting in the fall of 2010, a number of the papers reporting the Duke results have been retracted.

Over the time since the trials were suspended, there have been several significant developments. NCI asked the Institute of Medicine to develop principles for evaluating omics-based tests, and IOM released its report in 2012 ([IOM, 2012](#)). Drawing on lessons from the Duke case and informed by the development of other omics-based tests, the IOM report lays out a recommended development and evaluation process for these tests, and makes specific implementation recommendations to researchers, institutions, agencies, and journals ([IOM, 2012](#)).

Duke University has also taken steps to respond ([Califf, 2012](#)). Its Translational Medicine Quality Framework emphasizes new science and management approaches to ensure data provenance and integrity, the incorporation of adequate quantitative expertise, explicit management accountability in the institution beyond the individual lab for research affecting patient care, and enhanced conflict-of-interest reviews.

In 2015, ORI concluded that Potti had “engaged in research misconduct by including false research data,” citing specific examples of Potti's data that had been reversed, switched, or changed in a number of (now retracted) articles and other submissions ([ORI, 2015](#)). While Potti did not “admit nor deny ORI's findings of research misconduct,” he has expressed that he has no intention of applying for PHS (Public Health Service)–funded research, but agreed that if he is engaged with any PHS-funded research in the future, his research will be supervised for 5 years ([ORI, 2015](#)).

In this case, just about all the scientific checks and balances intended to uncover incorrect or fabricated research and protect human subjects failed over the course of several years. A summary of these failings illustrates some of the U.S. research enterprise's key vulnerabilities regarding integrity. Effective steps on the part of Duke to address the problems with Potti's work and investigate possible misconduct were delayed for years, and were finally triggered only by the disclosure of Potti's resume falsification. Those pointing out these problems were appropriately cautious about making formal allegations of misconduct, since there was a possibility that the problems were due to error or extreme sloppiness rather than falsification. Another contributing factor was the willingness of Joseph Nevins, a highly prestigious researcher, to vouch for the work and advocate for Potti with university administrators and others.

Individual Researchers

Anil Potti's misbehavior is at the center of the case. Prior to ORI's conclusion of research misconduct, Joseph Nevins and Robert Califf had both said that it is highly likely that Potti intentionally fabricated or falsified data ([CBS News, 2012](#)). In addition, Baggerly, Coombes, and Wang had documented many instances of sloppy or careless data analysis, and Perez documented use of unreliable predictors and omission of data not showing desired results. The negative impact of such sloppy and careless practices on the ability to replicate results and ultimately on patient care might be similar to the impact of fabrication or falsification.

In addition to problems with data and analysis, the IOM committee described a number of poor practices related to the clinical trials for the tests, including trials being undertaken simultaneously with preliminary studies ([IOM, 2012](#)).

Potti's collaborators also share responsibility. For example, despite being principal investigator of the lab where the research was undertaken, as well as Potti's mentor and coauthor, Joseph Nevins did not thoroughly check the original data files until after it was revealed that Potti had exaggerated his credentials in July 2010, more than 3 years after the data issues were originally raised ([CBS News, 2012](#)). Moreover, we now know from a deposition cited in court documents that Nevins “pleaded with Perez not to send a letter about his concerns to the Howard Hughes Medical Institute, which was supporting him, because it would trigger an investigation at Duke” ([Kaiser, 2015](#)). Indeed, Duke administrators testified to the IOM that none of Potti's coauthors (a total of 162 for 40 papers) raised any questions or concerns about the papers or tests until they were contacted by Duke at the start of the process of determining which papers should be retracted ([IOM, 2012](#)). Bradford Perez, the medical student described above, did raise concerns and removed his name from the papers that he contributed to, so his documented concerns were apparently not considered when that statement was made. Nevins remained on faculty as a department chair until his retirement in 2013, the year after the IOM report was released.

Institutional Policies and Procedures

In addition to the failures of individual researchers, lessons can be drawn from the responses by Duke as an institution during the controversy. Institutional shortcomings in policies and procedures, structure, systems, and oversight contributed to delays in recognizing that the science underlying the Nevins-Potti research was unsound. First, Duke's Institute for Genomic Science and Policy and its component Center for Applied Genomics and Technology, where Nevins and Potti worked, instituted its own system for undertaking clinical trials, separate from the extensive existing infrastructure of the Duke Cancer Center ([IOM, 2012](#)). This parallel pathway lacked the normal checks and balances as well as clear lines of authority and oversight.

In addition, systems for managing conflicts of interest at the individual and institutional levels were inadequate ([IOM, 2012](#)). For example, the IOM committee found evidence that researchers involved with undertaking the clinical trials had unreported financial or professional conflicts of interest. Some investigators held patents on one or more of the tests, or had links with one of the companies founded to market the tests. The institution itself, through its licensing relationships, had a financial interest in the success of the tests, as well as a reputational interest in having generated such an important new technology. It is of note that the institution had created a set of video and print materials featuring the research ([CBS News, 2012](#); [Singer, 2010](#)).

As noted in the 2012 IOM report, as a “responsible party” for assuring the integrity of the science conducted under their auspices, universities have particularly important responsibilities. These include responsibility for the hiring and promotion of the faculty members conducting research, the establishment and maintenance of oversight structures, and responsibilities for properly responding to and resolving questions about the validity of research or allegations of misconduct when they arise. It also includes the responsibility for ensuring the existence of an organizational culture and climate that sets expectations for research integrity that “are transmitted by the institution and modeled by its leadership. Institutional culture starts with the dean, senior leaders, and members of their team stating how research is to be conducted, with integrity and transparency, and with clarity that shortcuts will not be tolerated and that dishonesty is the basis for dismissal” ([IOM, 2012](#)).

The evidence now available, some that has come to light only after Freedom of Information Act requests and court depositions, suggests that Duke University and its leadership failed in virtually all of these responsibilities: for undertaking clinical studies outside the established review structures; for the failure to pursue internal investigation of serious, documented concerns until forced by outside forces to do so; for withholding from an external committee the full Baggerly/Coombes critique; for referring responsibility for

rechecking Potti's work back to the laboratory of his (explicitly conflicted) principal investigator, Joseph Nevins; for failing to employ the full set of institutional checks and balances that were in place; and for either incomplete or factually unsupportable statements made to the IOM Committee charged with examining the issue. The breadth and depth of these institutional failings are disappointing. Occurring in an institution of Duke's stature and resources, they raise troubling questions about the ability of research institutions, without more support and reinforcement, to manage complex cases when directed against prominent institutional researchers.

Duke suspended the trials and launched an investigation in the fall of 2009 in response to NCI concerns. However, this investigation had several serious flaws. Although the trials were resumed based on the report of the two external statistical experts, as noted above, these experts were not provided with several critical pieces of information. The IOM report also raises the possibility that Nevins was improperly in direct contact with the reviewers during the inquiry ([IOM, 2012](#)). As for the clinical trials that were undertaken based on the fabricated work, 117 patients were ultimately enrolled. Duke later faced a lawsuit brought by the families of eight of these patients, which was settled in May 2015. The terms of the settlement were not disclosed ([Ramkumar, 2015](#)).

In its Translational Medicine Quality Framework activity, Duke also identified an environment that might discourage postdocs or grad students from raising concerns with research within the lab or taking their concerns to others at the university as a possible problem. The university reported that it has established an ombudsman's office and taken other steps to address this.

Taken together, these institutional failings raise the question of whether, in addition to strengthening policies and procedures to the extent possible, research institutions should explore new mechanisms for bringing in outside perspectives in cases where it might be difficult for an institution to objectively address allegations of misconduct or other challenges to the soundness of science. In 2016, four members of the IOM committee published a piece critical of how Duke handled the case as an institution ([DeMets et al., 2016](#)).

Journal Policies and Practices

Although *Nature Medicine* and the *Journal of Clinical Oncology* did publish letters from Baggerly, Coombes, and Wang questioning the validity of data, along with responses from Potti, they rejected further questioning of the Duke results. This is likely the result of the common journal practice of not publishing additional comments on an article that appear to repeat concerns already raised in a previously published comment, so as to avoid involving the journal in an ongoing dispute. Further, other journals that had published other articles reporting the Nevins-Potti work were not responsive to questions raised by Baggerly and Coombes. This stance contributed to delays in recognizing the nature and extent of the problems with the papers. The translational omics case raises issues of how scholarly publishers, institutions, and the broader community should respond when the work underlying numerous papers in a variety of journals is questioned.

Sponsor and Regulator Policies and Practices

The IOM report identifies some ambiguities in Food and Drug Administration requirements for launching clinical trials on diagnostics as possibly contributing to the clinical trials being launched prematurely and to delays in finally shutting them down ([IOM, 2012](#)). The IOM report also points out that NCI felt constrained in communicating what it knew and the extent of its concerns with Duke and others early in the case, particularly before officials were aware that the agency was supporting aspects of the clinical trials ([IOM, 2012](#)). More direct and complete communication would be helpful in future cases.

THE RIKEN-STAP CASE

Synopsis and Rationale for Inclusion: *The RIKEN-STAP case illustrates issues that may arise related to authorship roles, mentoring, and data falsification. The extent to which coauthors should be held responsible for the data and findings of papers on which they are listed is a recurring question in many research misconduct cases.*

Yoshiki Sasai, a stem cell biologist of Japan's RIKEN research institute, committed suicide in August 2014 after the lead author on papers that he coauthored, Haruko Obokata, was found guilty of research misconduct ([RIKEN, 2014](#)). Obokata claimed to have found that a process that reprogrammed somatic cells into pluripotent cells by exposing the cells to stress; the authors termed the process “stimulus-triggered acquisition of pluripotency (STAP)” ([Obokata et al., 2014a](#), retracted). Obokata collaborated with Charles Vacanti's laboratory at Brigham and Women's Hospital, where the idea of STAP had supposedly originated ([Knoepfler, 2015](#)). Vacanti, professor of anesthesiology at Harvard Medical School and former chairman of the Department of Anesthesia at the Brigham and Women's Hospital, was a corresponding author on one of the papers, a coauthor on the other, and Obokata's mentor while she worked as a postdoctoral research fellow at Brigham and Women's Hospital.

Shortly after Obokata's findings were published in *Nature*, outside researchers were unable to replicate the study or achieve similar results, prompting an internal RIKEN investigation. The investigation committee concluded that she had fabricated data in at least one of the papers ([RIKEN, 2014](#)). The committee found problems with the data underlying the other papers, but was not able to conclude that fabrication or falsification had occurred because they did not have access to the original data ([RIKEN, 2014](#)). The committee found that Sasai had no involvement with the data fabrication, but bore a “heavy responsibility” for the incident because he did not insist that experiments be repeated even after problems with the data became obvious ([RIKEN, 2014](#)).

Both Sasai and Obokata made public apologies, but maintained that STAP works. Already disgraced, the Japanese media soon began to make “unsubstantiated claims about [Sasai's] motivations” and personal life, as well as shame him for a lack of oversight responsibility, all of which, Sasai wrote in a suicide note, drove him to take his own life ([Cyranoski, 2014c](#)). Vacanti also maintained “absolute confidence” in the phenomenon and released follow-up protocols to the retracted *Nature* papers to assist in the reproducibility of STAP cells ([Vacanti and Kojima, 2014](#)). Following RIKEN's investigation and the retraction of the *Nature* papers, Vacanti stepped down as chairman of the Department of Anesthesia at the Brigham and Women's Hospital and took a 1-year sabbatical from his professorship at Harvard Medical School. He did not reference the STAP case in his letter of resignation from Brigham and Women's Hospital.