Sci Eng Ethics (2019) 25:621–629
https://doi.org/10.1007/s11948-018-0027-3

COMMENTARIES

Correctable Myths About Research Misconduct in the Biomedical Sciences

Barbara K. Redman[1]

Received: 28 October 2017 / Accepted: 29 January 2018 / Published online: 5 February 2018

Abstract A recent National Academy report on research integrity noted that policies are not evidence-based, with no formal entity responsible to attend to this deficit. Here we describe four areas of research misconduct (RM) regulations governing Public Health Service funded research that are empirically and/or ethically questionable. Policies for human subject protection, RM and conflict of interest are not harmonized, making it extremely difficult to deal with complex cases which often contain allegations in all of these areas. Second, detection of RM has depended entirely on whistleblowers in spite of evidence of significant under-reporting. Third, the scientific record is far from cleansed of the effects of falsified/fabricated work through current mechanisms of retraction. Finally, lack of fairness in the regulations may reflect lack of a Belmont Report-like document to guide ethics of RM policy. These issues are likely common in other countries. RM regulations should be harmonized with related regulations and their effectiveness tracked, open access to data for independent replication and improved statistical tests are an essential supplement to whistleblowers, correction of the scientific record will require a major effort, and further ethical analysis and guidance are as important as is empirical study for the improvement of RM regulations. Further consideration should be given to assigning current regulations for human subjects protection, RM and conflict of interest to a single authority and to the further development of a Belmont-like report of essential principles, for RM.

Keywords Research misconduct · Research ethics · Whistleblowers · Research regulation

✉ Barbara K. Redman
br68@nyu.edu

[1] Division of Medical Ethics, New York University Langone Medical Center, 227 East 30th Street, #753, New York, NY 10016, USA

Introduction

Much of the federal regulation undergirding areas of research integrity has not been evidence-based. No formal entity has as its responsibility research on the effectiveness and unintended consequences of existing research policy or review of proposed new policies (NASEM 2017). Lacking such evidence, policy is likely to be based on best judgment and negotiation with the science community. Adding to the urgency of seeking better options for preventing, detecting and managing RM is the recent National Academy of Sciences panel conclusion that "the research enterprise faces serious challenges in creating appropriate conditions to foster and sustain the highest standards of integrity" (NASEM 2017, p 2). The goal is a level of fidelity in which in a cumulative process science produces a body of reliable knowledge (NASEM 2017).

Here we consider four areas of policy on research misconduct (RM), governing the Public Health Service (PHS) funded research and implemented by the US Office of Research Integrity (ORI), that are now ethically and/or empirically questionable. While embedded in current federal policy, they can be considered myths, based on legend—a traditional story about what should work—but not supported by evidence or by accumulated experience that now shows them to be problematic. Myths were identified through an extensive review of the literature on management of RM, and were thought to be foundational to current regulatory arrangements. While there may be additional myths, examination of the four is meant to demonstrate fragile assumptions that should be the basis for examining current policy.

The paper first identifies the four myths, followed by a discussion of each including options for resolution. Experiences of other nations support a search for effective strategies that are evidence-based. Discussion suggests ways in which RM policy might be updated, to address the myths including a single agency.

Some Myths About Research Misconduct Policy

Myth #1: Three active areas of federal policy—human subjects protection, prevention and detection of research misconduct and control of conflict of interest—can be adequately governed by separate regulations and administered by different agencies as is currently the case.

Myth #2: Whistleblowers as monitors for RM are sufficient. The trigger for detection is notice by a complainant, defined in federal regulations as "a person who in good faith makes an allegation of research misconduct." (42 CFR 93.210). Much research shows the vulnerability of such persons in all organizational settings and how few actually lodge a complaint.

Myth #3: RM policy corrects the scientific record and restores it to accuracy. "In response to a research misconduct proceeding, HHS may impose HHS

administrative actions that include but are not limited to (1) clarification, corrections or retraction of the research record." (42 CFR 93.407(a)(1)

Myth #4: Current RM policy is fair to all parties involved. The governance of RM should restrict the harm from it to all relevant parties and should be fairly applied.

Myth #1: Separate Regulations

The purpose of research misconduct regulations is to "protect the health and safety of the public, promote the integrity of PHS supported research and the research process, and conserve public funds." (42CFR 93.101(e) The RM regulation does not refer to the regulation for the protection of human subjects (CFR 45 Part 46). Indeed, RM is not limited to studies involving human subjects. But when there are allegations that more than one set of regulations have been violated, the lack of harmonization among them is problematic.

The IRB/RIO/IO Working Group addressed this issue and published their report in 2014. They noted that these differences were believed to be significant and fundamental, making simultaneous compliance with both sets of standards for the same case unnecessarily complicated. Human subjects protection and RM regulations involve different oversight officials and procedures, standards of proof, expectations of privacy and appeal procedures, making it very difficult to manage cases with allegations in both areas (Bierer et al. 2014). Addition of unclear conflict of interest standards adds another layer of confusion.

Complex research ethics cases almost always involve violations in several regulated domains. A recent example from Duke University involved research in development of genomic predictors for cancer risk, assessment and treatment. The investigator, Anil Potti, was found to have committed RM by falsifying data, the IRB did not recognize that a genomic test is a medical device subject to FDA requirements, and the university did not respond to multiple whistleblowers, thus failing in its oversight responsibility, arguably part of reputational conflict of interest (De Mets et al. 2017).

A single authority managing the several areas of research regulation would be preferable to the historical accident of creating a new agency for each additional area of regulation and through harmonization of policy and procedures, decrease the institutional regulatory burden. The current move to the conceptual approach of research integrity signals an understanding that science is produced in an evolving system of individuals, organizations and policies, each of which interacts with and affects the functioning of the others. Reflecting this interdependence, consider one action—falsifying/fabricating research informed consent forms. Such an action simultaneously violates regulations for human subjects protection, RM and may be motivated by a conflict of interest, for which some regulations exist. How an institution is to proceed procedurally is one important question. A second is whether a unified analysis of all three regulated areas could assess the adequacy of the design and/or implementation of the oversight system which could have prevented/detected this violation as well as the causative factors such as lack of knowledge or pressure from perverse incentives, which should be ameliorated. Such analysis would be

one element, along with the quality of the science, for constant quality improvement toward research integrity.

Under current policy an IRB might detect and deal with a conflict of interest for a research proposal, if it had been divulged, and with scrupulous follow-up monitoring might note the improper informed consents. Under RM regulations a whistleblower would need to come forward. How frequently such diligence occurs is unknown, which in itself, signals a weak regulatory system.

Myth #2: Whistleblowers (Complainants) as Prime Detectors of RM

If a complainant's allegation involves PHS supported research or research training, fits the definition of fabrication, falsification or plagiarism in proposing, performing, or reviewing research or in reporting research results (42 CFR 93.103) (FFP) and is sufficiently credible and specific so that potential evidence may be identified (42 CFR 93.307), an inquiry can be initiated by the institution receiving the funds. The party making an allegation can be an editor, a reader or an outside researcher but more frequently is a colleague or other participant in the research environment (Stroebe et al. 2012).

A 2008 survey of NIH funded researchers asking if they had observed or had direct evidence of RM in their own department revealed a dramatic under-reporting of this evidence to institutional or regulatory authorities (Titus et al. 2008). Even though it strikes at the core of the regulatory strategy of using whistleblowers as the prime guardian of research conduct, this finding should not be surprising. Moral identity, licensing and disengagement shape moral behavior (Redman and Caplan 2017). Group norms often become an anchor for good or bad rules and behaviors, provide guidance in uncertain situation, enforce norms and sometimes blind one's internal moral compass (Ellemers 2017). Studies of RM have rarely identified these influences but instead have focused almost entirely on the reluctance of whistleblowers to come forward. Additionally, whistleblowing is very unlikely to occur if the organization considers acts of FFP normal and thus futile to report.

New scientific practices and technologies can supplement infrequent allegations from whistleblowers. Some editors and funders require the deposition of raw data and code from a study, available for replication which can expose RM. Systematic reviews could detect evidence of potential RM but infrequently do so (Elia et al. 2016). Technologies to document plagiarism are now widely used. Multiple statistical approaches to uncover data anomalies which may be RM, are being perfected (Carlisle and Loadsman 2017; Simonsohn 2013; Geyer and Williamson 2004) and after additional development should be widely used. Fraud monitoring and recovery plans are increasingly required for clinical trials to catch falsification of eligibility criteria or patient diaries or under-reporting of adverse events (Herson 2016).

For prevention and detection of misconduct, other fields have adopted the fraud triangle: opportunity to cheat, motivation to cheat, ability to rationalize or justify dishonest behavior (Houdek 2017). A preventive strategy requires setting and enforcing standards including through inspections, identifying individuals who believe they are "in a hole" with dishonest behavior being the only way out, and the

ability to rationalize that what they are doing is justified and thus moral. If all three elements of the fraud triangle are present, the risk is thought to be higher (Houdek 2017). Current RM regulations do not address these issues but rather leave them nebulously to the "university environment" without attending to whether the work environment may contribute to all or some elements of the fraud triangle.

Myth #3: Restoration of the Scientific Record

This is an old myth, built on the assumption that the scientific record is largely correct. Recent research has found that much published medical and other research is not reliable or its reliability is uncertain or is not replicable (Ioannidis et al. 2017).

The mechanism of correction when results are no longer considered to be valid, whether from pervasive errors or RM, is retraction. Unfortunately, this system is flawed in several ways. There is a lack of standards or consistency in issuing retraction notices. Many represent the authors' view of events, and notices may not be prominently published if at all. Misconduct may be disguised as error or irreproducibility. Not all articles known to be erroneous are retracted and the actual number of retractable articles in the literature is unknown (Casadevall et al. 2014).

Given this lack of rigor in mechanisms used to correct the scientific literature, what can RM policy do to adopt rigorous standards? First, many cases closed by ORI require retraction of publications in which evidence of F/F has been found. Resnik and Dinse (2013) found that the 208 closed cases from 1992 to 2011 yielded 174 articles to be retracted, of which 127 were found, which means that 15% of potential retraction notices were likely never published.

Retracted work may continue to be cited and used as justification for treatment and/or for new studies, placing additional subjects and patients at risk. For example, the 1988 case of Stephen Breuning, whose claim that psycho-stimulant drugs were more effective than neuroleptic drugs for what were then called "retarded" children, was reported in 20 published papers. Eight were discredited by RM reviewers, with three retracted. Twenty-four years later, some affirmative cites of this work continued (Korpela 2010). IRBs, funders and journal editors could request evidence that retracted material is not incorporated in a proposal. CrossMark provides information about corrections and retractions but is far from universal.

Myth #4: Fairness of the Regulatory System for RM

As originally enacted RM regulations were meant to identify those who fabricate, falsify or plagiarize research and to remove them from the scientific community. In the 20 + years since they were adopted, additional perspectives and evidence about fairness have come to light.

The situation in which individual scientists operate is open to large numbers of incentives that do not support ethical science and which scientists themselves do not control. In the Duke case noted earlier, De Mets et al. (2017) questioned why only the investigator was found to have committed RM but the University was not charged for lack of proper oversight.

Some WB are acting out of bad faith, making allegations of RM solely to protect a commercial interest or to stop a competitor's program of research through prolonged investigation. Such a tactic, common in chemical, tobacco, petroleum and other industries, was lodged by an individual from the lead industry against pediatrician Herbert Needleman, who had found that lead levels much lower than then-current standards were damaging to children. Although eventually cleared, Needleman suffered considerable disruption to his research and resources to defend himself (Watts 2017). While RM regulations specify that allegations must be made in good faith, current regulatory systems seem unable to identify and control bad faith whistleblowers, perhaps because conflict of interest policies are so weak and not incorporated into RM policies.

In addition to evidence that the scientific record is incompletely corrected, studies do not track the degree of harm if any to research subjects in the index or followup trials containing or building on tainted data or to subsequent patients, who are treated based on fabricated/falsified data.

Finally, the under detection of RM through whistleblowers means that similar infractions are not treated similarly, open to the vicissitudes of WB intentions and motivations. In addition questions have been raised whether sanctions to individuals found to have committed RM are proportional to the infraction and on what basis—strength of the level of intent or recklessness, amount of FFP or consequences from it or some combination of these factors? No such algorithm has been made public. Many of those found to have committed RM are excluded from contracting with the US government and serving on PHS advisory committees for periods varying from 3 to 7 years (ORI website), suggesting that the length of the sanction is proportionate in some way.

Unlike the human research subjects regulatory system built on the ethical analysis contained in the Belmont Report, the regulatory system for RM does not have such an analytical framework, which could help to detect and resolve issues including fairness. Extension of Belmont principles—respect for persons, beneficence and justice (Belmont 1979)—applied directly to RM would clearly identify the importance of harm from fabrication/falsification to research subjects as well as violation of respect for persons. RM investigations rarely fully identify and quantify harm but should do so in a unified approach to research integrity.

In support of research integrity, the National Academy report identified central scientific values such as honesty, accountability, fairness, objectivity and stewardship (NASEM 2017). All are relevant for RM but should be developed into a more explicit statement about how they apply. Of particular interest are violations which occur in use of research methods, long considered to be the self-regulated domain of science and not of public interest as is protection of research subjects. There is some evidence of uncertainties on the part of scientists about methodological rules and standards and a need for much more clarity (Hangel and Schickore 2017). This could indicate that scientists would not agree that certain methodological practices constituted fabrication/falsification.

The social institution of the IRB helps scientists meet requirements for protection of human subjects (Douglas 2014); there is no parallel institution for RM. Developing one would require resolution of uncertainties where they exist. Currently, there

is some piecemeal attention to methodological shortcomings of various scientific fields as they happen to be exposed or not. A body parallel to the IRB could be developed separately or preferably incorporated into an expanded institution encompassing all of the elements of research integrity.

Are These Largely US Issues?

Research misconduct is currently directed by national regulatory/self-regulatory agreements, with varying balances in countries. Bosch's 2010 review of European country RM policy finds that the Scandinavian countries were first to draw up regulations. A current review carried out by Horizon 2020 found no consensus across Europe on the meaning of RM or RI and was largely dependent on self-regulation and not on legislative instruments (Fuster and Gutwirth 2016). Arrangements vary. In the UK there are no statutory frameworks to regulate research integrity and sanction RM apart from medical self-regulation supported by the state. Examination of the General Medical Council's fitness to practice cases revealed that RM frequently did not fit that model with no consistent sanctions if any at all (Jacob 2016). Many European countries, both regulated and not, have experienced RM scandal cases reported in the scientific press (Netherlands, Germany, Italy, Sweden, Norway, UK).

In Asia, the South Korean case of Woo Suk Hwang prompted significant regulatory reforms in subject protection, RM, financial management and conflict of interest. After cases of massive fraud, Japan has adopted regulations. In the face of a serious problem, China has adopted strong punishment for faked clinical trial data (Cyranoski 2017). Less is known about the incidence of RM or policies to address it in Latin America or Africa.

It is reasonable to think that the myths outlined above are relevant to many countries seeking to detect and manage RM.

Discussion

Regulatory systems must be updated regularly to address new technologies, empirical information and altered social arrangements. For the first two of these myths, potential or actual remedies have emerged—coordination of policy to protect subjects and positioning the scientific community rather than vulnerable whistleblowers to detect RM in part through increased transparency and use of additional statistical tools. While cleansing the scientific literature is a much larger project, proper implementation of retraction and citations from RM would be helpful.

The ethics of regulating RM have not been well thought out. Harm is rarely addressed, thus lessening the moral sting of confronting damage done which is sometimes irreversible to an individual's health from the application of falsified research findings. Just regulation requires much more attention to organizational and system roots of individual misconduct; currently, organizations can just search out the 'bad apple' and consider the job done.

Ascendance of the integrating construct of research integrity in the US (NASEM 2017) and Europe through Horizon 2020 (Fuster and Girwith 2016) focuses on the production of reliable research in as efficient a manner as possible while protecting subjects. It requires simultaneous attention to all the elements of the research system producing knowledge including prevention, detection and reversal of ethical and methodological breakdowns. Current regulatory domains (human subjects protection, RM, conflict of interest) are inevitably highly inter-related. Managing them separately with weak oversight neglects effective diagnosis and amelioration of causes and correction in all relevant elements of the research production system.

Further explication of ethical principles undergirding management of RM should include the extension of Belmont principles developed for human subjects protection to RM as well as additional foundational work on methodological norms and standards, now apparently not well clarified. An institution parallel to the IRB with explicit norms and standards for RM could eventually be folded into a single authority encompassing all of the elements of research integrity.

Questions raised and experience with the four myths identified in this paper suggest ways in which current regulations no longer provide sufficient guidance. To build and sustain trust in the regulatory system, new options should be considered.

References

Belmont Report, Federal Register 1979 April 18; 44(76), 23192–23197.

Bierer, B., Mark, B., & on behalf of the IRB/RIO/IO Working Group. (2014). Research misconduct involving noncompliance in human subjects research supported by the Public Health Service: Reconciling separate regulatory systems. *The Hastings Center Report, 44*(4 Special Number), S2–S26.

Bosch, X. (2010). Safeguarding good scientific practice in Europe. *EMBO Reports, 11*(4), 252–257.

Carlisle, J. B., & Loadsman, J. A. (2017). Evidence for non-random sampling in randomized controlled trials by Yuhji Saitoh. *Anaesthesia, 72,* 17–27.

Casadevall, A., Steen, R. G., & Fang, F. (2014). Sources of error in the retracted literature. *FASEB Journal, 28*(4), 3847–3855.

Cyranoski, D. (2017). China cracks down on fake data in drug trials. *Nature, 545,* 275.

De Mets, D. L., Fleming, T. R., Geller, G., & Ransohoff, D. F. (2017). Institutional responsibility and the flawed genomic biomarkers at Duke University: A missed opportunity for transparency and accountability. *Science and Engineering Ethics, 23,* 1199–1205.

Douglas, H. (2014). The moral terrain of science. *Erkenn, 79,* 961–979.

Elia, N., et al. (2016). How do authors of systematic reviews deal with research malpractice and misconduct in original studies? A cross-sectional analysis of systematic reviews and survey of their authors. *British Medical Journal Open, 6,* e010442.

Ellemers, N. (2017). *Morality and the regulation of social behaviors; groups as moral anchors*. London: Routledge.

Fuster, G., & Gutwirth, S. (2016). *Promoting integrity as an integral dimension of excellence in research*. D II.4 Legal analysis, 7/6/2016. http://www.hioaresponsibleinnovation.wordpress.com/projects/printeger. Funded from European Union's Horizon 2020 research and innovation programme.

Geyer, C. L., & Williamson, P. P. (2004). Detecting fraud in data sets using Benford's Law. *Communications in Statistics, 33*(1), 229–246.

Hangel, N., & Schickore, J. (2017). Scientists' conceptions of good research practice. *Perspectives on Science, 25*(6), 766–791.

Herson, J. (2016). Strategies for dealing with fraud in clinical trials. *International Journal of Clinical Oncology, 21,* 22–27.

Houdek, P. (2017). Professional identity and dishonest behavior. *Society, 54,* 253–260.

Ioannidis, J. P. A., Stuart, M. E., Brownlee, S. S., & Sherri, A. (2017). How to survive the medical information mess. *European Journal of Clinical Investigation, 47,* 795–802.

Jacob, M. (2016). On the scope and typology of research misconduct: The gaze of the General Medical Council, 1990–2015. *Medical Law Review, 24*(4), 497–517.

Korpela, K. M. (2010). How long does it take for the scientific literature to purge itself of fraudulent material: The Breuning case revisited. *Current Medical Research and Opinion, 26*(4), 843–847.

National Academy of Science, Engineering and Medicine. (2017). *Optimizing the nation's investment in academic research*. Washington, DC: National Academy Press.

Redman, B., & Caplan, A. (2017). Improving research misconduct policies. *EMBO Reports, 18*(4), 511–514.

Resnik, D. B., & Dinse, G. E. (2013). Scientific retractions and corrections related to misconduct findings. *Journal of Medical Ethics, 39*(1), 46–50.

Simonsohn, U. (2013). Just post it: The lesson from two cases of fabricated data detected by statistics alone. *Psychological Science, 24*(10), 1875–1888.

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on Psychological Science, 7*(6), 670–688.

Titus, S. L., Wells, J. A., & Rhoades, L. J. (2008). Repairing research integrity. *Nature, 453,* 980–982.

Watts, G. (2017). *Herbert Leroy Needleman, obituary, Lancet 930*.