Reflections on the Irreproducibility of Scientific Papers

Roberto Bolli

This is the first of two back-to-back editorials1 in which I discuss two different, but related, topics: the irreproducibility of research work and the importance of the test of time, respectively. These topics are related because, as we will see, only with time can one tell whether a scientific study is reproducible, and because reproducibility would increase dramatically if the value of research work were gauged on the basis of the test of time rather than its short-term impact.

The lack of reproducibility of scientific papers is an enormous problem that plagues biomedical research and seems to be getting worse. It undermines the entire research enterprise, for reproducibility is the very essence of science; without reproducibility, there is no science.

Because of the importance and magnitude of this problem, the News and Views article in the current issue of the journal2 is devoted to a discussion of the causes of, and potential remedies for, irreproducibility. The article, written by Ruth Williams, distills her discussions with several editors of Circulation Research and other scientific leaders. I hope the readers will find this article thought-provoking and conducive to a constructive discussion of how to improve the situation.

Irreproducibility is devastating because it paralyzes scientific progress, causes fictitious “controversies,” and undermines the very credibility of science. A good example is in the field of myocardial infarct size limitation (a.k.a. cardioprotection). Although >20,000 papers have claimed to have found a therapy that reduces infarct size, rarely have these results been reproducible from one laboratory to another. This has been the major reason why the field of cardioprotection has not been translated into any clinical therapy despite 40 years of intense work and billions of taxpayer dollars spent by the NIH and other bodies to find a drug that limits myocardial infarct size. I am certain that readers can think of many other egregious examples of irreproducibility.

What are the causes of this scourge? Many of them are discussed in the accompanying News & Views article.2 Here, I present briefly my opinions on the subject (Table).

In many cases, the inability to reproduce someone else’s results stems from differences in experimental settings or methods. For example, different strains of mice respond differently to myocardial ischemia and reperfusion; consequently, it is not surprising that different laboratories arrive at different conclusions if they use different mouse strains, even if everything else is done in an identical manner. Likewise, studies of left ventricular remodeling in the setting of a permanent coronary artery occlusion cannot be expected to be reproduced in the setting of a temporary coronary occlusion and vice versa.

An important reason for irreproducibility is that most papers do not provide sufficient details in the Methods section for other investigators to reproduce the work that is described. Sometimes this is due to page limitations; sometimes to failure to appreciate the importance of details; sometimes to the authors’ conscious decision to withhold key information in hopes of keeping their competition at bay; sometimes to plain laziness. Whatever the reason, the Methods section is frequently incomplete, superficial, and/or vague, making it difficult to know how a study was actually performed. Details that are crucial to reproduce a result are often not mentioned. Most of us know of some specific technical minutia (sometimes referred to as a “trick”) that is critical for the experiment to succeed; yet, rarely are these “tricks” described in published articles. Particularly annoying is the practice of some authors to cite a previous paper, which in turn cites a previous paper, and so on until one discovers that the original paper does not actually describe the method that it is supposed to describe. It is my opinion that all published manuscripts should include a full description of the methodology employed, so that others can reproduce those results. Accordingly, at Circulation Research, the editors request that each manuscript be accompanied by a Methods section (published online) that must be sufficiently detailed and specific to enable others to reproduce the experiments. This policy is stated in the Instructions to Authors, and authors are urged to follow it meticulously.

Other reasons for lack of reproducibility include insufficient technical prowess, sloppiness in the design and/or execution of the experiments (small group sizes, badly performed assays, lack of attention to confounding variables, poor methodology, etc), and subtle variations in the protocol that may seem inconsequential but actually have a major impact on the results of the study. Again, one cannot overemphasize the concept that, in research, small details can make a huge difference. For example, just changing the source or even the lot number of a reagent can make the data irreproducible. Accordingly, when submitting papers to Circulation Research, authors should include in their online Methods section all the details that can affect the outcome of a study, such as sources of reagents and lot numbers of antibodies.

The most dreaded cause of irreproducibility is, of course, scientific fraud. The actual prevalence of fraud in research is unknown but is likely to be underestimated. Most instances of fraud do not consist of outright fabrication of data (which I believe is rare), but rather of subtler manipulations, such as selective data exclusion (not prospectively declared), biased assignment of animals to experimental groups, use of subjective measurements,
Science, in particular, it appears that few people, if any, care. It seems that things are moving too fast for us to stop and think. When I was a young investigator, it seemed to me that lack of reproducibility was a big deal; now, their careers depend, at least in part, on the reproducibility of their work, they would be more careful with what they publish. In most cases, the fact that a study cannot be reproduced generates little, if any, interest because every one is focused on the new papers that appear on a daily basis. There is just not enough time to do that when one is trying to keep up with the torrent of new information that comes out every day.

The result is that authors are rarely held responsible for what they publish. In most cases, the fact that a study cannot be reproduced generates little, if any, interest because everyone is focused on the new papers that appear on a daily basis. It is likely that this awareness encourages some investigators to publish sloppy or even fraudulent work, since they know that, in all probability, there will be no major consequences.

Countering these trends will be arduous, if not impossible, for it will require a change of culture, whereby the focus of the scientific community shifts from the short term to the long term and investigators are rewarded not for "catchy," "sexy," or "hot" work, but for valid work; that is, work that is thorough and rigorous, and, therefore, reproducible. Such a change in the reward mechanism, if implemented at the level of, say, study sections and promotion committees, would encourage more careful work while discouraging sloppiness and fraud, neither of which is reproducible or can stand the test of time.

Unfortunately, I believe that a cultural change is not plausible, nor can any of us change contemporary culture. What editors can do instead is implement procedures that will alleviate the causes of irreproducibility. The editors of Circulation Research believe that this is an issue of utmost importance, and a number of initiatives are in the pipeline. Stay tuned.

### Disclosures

None.

### References

1. Bolli R. Dandum semper est tempus: The crucial importance of (and increasing disregard for) the test of time. Circ Res. Published in the October 9, 2015 issue. doi: 10.1161/CIRCRESAHA.115.307613.

### Table. Causes of Irreproducibility in Biomedical Research

| 1. Differences, at times subtle or inadvertent, in experimental settings, methods, and/or protocols |
| 2. Insufficient methodological details in published manuscripts (due to page limitations, laziness, failure to appreciate the importance of details, authors’ conscious decision to withhold key information in hopes of keeping their competition at bay, etc) |
| 3. Insufficient technical prowess |
| 4. Poor experimental design (eg, insufficient statistical power, poor choice of assays) |
| 5. Sloppy execution of the experiments (eg, deliberate deviations from protocol, lack of attention to confounding variables) |
| 6. Lack of rewards for reproducibility and disincentives for irreproducibility |
| 7. Fraud (prevalence likely underestimated; subtler forms of fraud likely to be more prevalent than commonly thought) |

arbitrary choice of data for quantitative analyses (eg, microscopic fields, echocardiographic images, Western immunoblots), etc. While we are all familiar with high-profile cases of fraud, I think that that is only the tip of the iceberg; for each case of fraud that is discovered, I suspect there are many more, subtler cases that elude detection. Less blatant (and thus unrecognized) forms of fraud are probably more common than most people think.

So, what can be done to improve the situation? One way in which editors and reviewers can increase reproducibility is to demand that authors provide a detailed description of their methods, as indicated above. Circulation Research has implemented this policy for more than five years. In addition, editors and reviewers should be sensitive to any indication of data manipulation and utilize available tools to exclude it or confirm it. While ruling out data manipulation in every submitted paper would be nearly impossible, it is realistically feasible if done in those papers that are being considered for publication. At Circulation Research, editors and reviewers are asked to scrutinize carefully all papers, particularly those that are headed toward acceptance. Although editors have the option of scrutinizing all figures of accepted papers with the use of software that aids in the detection of data manipulation, in practice this is an expensive and time-consuming remedy that few journals use.

In an ideal world, discordant studies should trigger an in-depth discussion of the possible reasons for the discrepancy, but given the increasing pace of research and the sheer volume of articles that appear every day, this is difficult in many cases. Nevertheless, I believe that authors should be held responsible for what they publish, and that a track record of papers that cannot be reproduced should be seen by peers as a red flag. This may discourage investigators from publishing premature, sloppy, and/or outright fraudulent data. Conversely, publication of reproducible papers should be valued, emphasized, and rewarded in such venues as grant review groups, promotion committees, prize committees, etc. At present, there is inadequate reward for publishing reproducible work and inadequate disincentive not to publish work that may or may not be reproducible. If investigators knew that their careers depend, at least in part, on the reproducibility of their work, they would be more careful with what they publish. I would be remiss here if I did not frame this issue in the broader context of our culture, which is certainly not conducive to introspection and in-depth analysis of any issue, including scientific disagreements. Specifically, I believe that the problem of irreproducibility is exacerbated by the short attention span of our culture. When I was a young investigator, it seemed to me that lack of reproducibility was a big deal; now, it appears that few people, if any, care. It seems that things are moving too fast for us to stop and think. Science, in particular, is moving faster and faster, at an increasingly hurried pace. Astonishingly, ≈1400 articles are published every day in biomedical fields. If I recall correctly, it has been estimated that the amount of new information generated every year greatly exceeds all of the information generated from the dawn of civilization to the year 2000. With the accelerated pace of work and publication, everything now moves so fast that the attention span of readers has become quite short and most papers are forgotten soon after they are published. Besides, there are so many new articles to read that few bother to figure out why a particular paper cannot be reproduced. There is just not enough time to do that when one is trying to keep up with the torrent of new information that comes out every day.

The result is that authors are rarely held responsible for what they publish. In most cases, the fact that a study cannot be reproduced generates little, if any, interest because everyone is focused on the new papers that appear on a daily basis. It is likely that this awareness encourages some investigators to publish sloppy or even fraudulent work, since they know that, in all probability, there will be no major consequences.

Countering these trends will be arduous, if not impossible, for it will require a change of culture, whereby the focus of the scientific community shifts from the short term to the long term and investigators are rewarded not for "catchy," "sexy," or "hot" work, but for valid work; that is, work that is thorough and rigorous, and, therefore, reproducible. Such a change in the reward mechanism, if implemented at the level of, say, study sections and promotion committees, would encourage more careful work while discouraging sloppiness and fraud, neither of which is reproducible or can stand the test of time.

Unfortunately, I believe that a cultural change is not plausible, nor can any of us change contemporary culture. What editors can do instead is implement procedures that will alleviate the causes of irreproducibility. The editors of Circulation Research believe that this is an issue of utmost importance, and a number of initiatives are in the pipeline. Stay tuned.
Reflections on the Irreproducibility of Scientific Papers
Roberto Bolli

Circ Res. 2015;117:665-666
doi: 10.1161/CIRCRESAHA.115.307496

Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/117/8/665

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org/subscriptions/