Perspective

Research Integrity

PETER BUTTRICK, MD

Aurora, CO

This past fall, a large number of foundational papers (>30) that described cardiomyocyte plasticity and laid the groundwork for clinical trials using both putative bone marrow derived and resident cardiac stem cells as a regenerative therapy were retracted. In parallel, a major research institution was fined $10 million to settle claims related to this work.[1] Concurrently, an NHLBI sponsored clinical trial (CONCERT_HF) that was designed to evaluate the effect of two different types of stem cells (mesenchymal and c-kit + cells, alone and in combination) in patients with ischemic heart failure was paused and will likely not restart.[2] More recently, a second prestigious research institution was fined more than $100 million and a decade's worth of work in the pulmonary hypertension/fibrosis realm was called into question. This is perhaps an unprecedented moment of scientific caution or even rebuke, what Jon Epstein has eloquently referred to as "a time to press reset".[3] As a heart failure research community, it is incumbent upon us to reflect on the cardiac regeneration story: how we got here and also on what precautions ought to be in place to prevent another similar episode in the future.

A bit of history: the concept that the heart had significant regenerative capacity was first proposed almost 20 years ago and conceptually was very exciting at the time and truly paradigm shifting.[4] However, the finding was controversial and several studies were published within short order questioning the premise,[5–7] and the research community became somewhat polarized between those who embraced the concept and those who were skeptical. Much like in today's political landscape, data were viewed subjectively and as a result, some of the supportive publications were not as critically reviewed as perhaps they should have been. Because of the perceived conceptual novelty and therapeutic potential and viewed in the context of a paucity of alternative therapies, there was a great deal of enthusiasm (and significant financial resources were mobilized) to move quickly into clinical trials. The objective data from these initial clinical studies (with relatively small numbers of subjects) were unimpressive, results that were felt to reflect a variety of confounders, including the cell types chosen for injection and the disease context. Meta-analyses of multiple clinical trials (using heterogeneous therapeutic strategies) showed at best a very modest effect on ejection fraction, which correlated with the number of cells injected and the timing of the injections (perhaps a confounder and not a direct biologic consequence) and no effect on clinically relevant but hard to achieve endpoints such as death and hospital readmission.[8] Despite this, enthusiasm for a regenerative strategy driven by cell injections of different immature cell populations continued and is still ongoing, with each negative or marginally positive result eliciting not a fundamental reassessment of the approach but rather a modification of the strategy. More recently, powerful scientific studies using contemporary techniques that are very difficult to refute have laid to rest the concept that cardiocyte proliferation is an intrinsic reparative mechanism in the heart and this has now been followed by the retractions cited above.[9,10]

Given this history, it is incumbent on us to ask several questions. First, why were the initial data sufficiently compelling that they essentially withstood a barrage of contradictory work? Second, why was the community so eager to move into clinical trials before a solid body of preclinical work could be assembled? Third, why was the clinical research community not more skeptical of what was essentially a large body of negative results? And of course the answers lead to a larger question namely what steps should be taken, or what restraints need to be in place, to prevent this history from repeating itself?

Viewed from the standpoint of a journal editor, the answer to the first question probably resides in our strong desire to publish high-impact work. Studies that are positive *and* paradigm-shifting certainly fall into this category. Publication bias is a real thing but this does not imply that it is wrong to encourage work that seeks to "push the envelope". Indeed the historical march of science has been defined by risk taking, or what Thomas Kuhn characterized as paradigm shifts. Balanced against this, is the need for rigor. The recent investigation of reproducibility in both social

From the University of Colorado, Aurora, CO.

Manuscript received April 4, 2019; revised manuscript accepted April 4, 2019.

Reprint requests: Peter Buttrick, Division of Cardiology, University of Colorado, University of Colorado Health Sciences Center, 12631 East 17th Street, M/C B130, PO Box 6511, Aurora, CO 80045. E-mail: peter.buttrick@ucdenver.edu

1071-9164/$ - see front matter

https://doi.org/10.1016/j.cardfail.2019.04.005

science and in translational biology has, with some exceptions, concluded that imperfect statistical analytics often underlie studies that don't reproduce, such that a contemporaneous rigorous analysis might well have identified flaws. Examples of this include underpowered study designs, selection models, and excess (or selective) significance determinations.[11–13] Beyond this, there are a number of factors that can be imposed by journals and journal editors to enhance rigor, such as study registration and prespecified analytical plans (in the case of clinical trials), and open materials and access to primary data for reanalysis.[14] Whether any of these tactics would have identified the errors in the now withdrawn body of work is hard to say; however, as a strategy going forward it is reasonable to demand that these issues are addressed and acknowledged especially when a piece of work has the potential to refocus the field.

In this context, a recent study evaluating reputation assessment of scientists is quite compelling.[15] This study demonstrated that both peer and public assessments of scientists were based more on how they pursued knowledge, not whether their initial results were true. When comparing a scientist that produced boring but certain results with another that produced exciting but uncertain results, opinion favored the former despite the belief that more rewards accrued to the latter.

The second question addressing why and how the flawed basic observations were accelerated into clinical trials is a similarly complex one. Health-care innovation is both an expensive and a lucrative area and there are incentives both for "inventors" and tech start-ups to be first in class with a new technology or biologic strategy. With a potential market, such as novel heart failure therapy, that is in the billions of dollars, these incentives are undoubtedly amplified. Superimposed on this are patient preferences and also institutional needs to be the innovative center within a commercial market in order to magnify visibility and by extension, market share. The rush to advance this plausibly unproven approach was likely driven by a number of these factors and it is hard to know how to impose some checks on this tendency going forward. In defense of the spate of clinical trials that were launched, aside from small sample sizes that were dictated by cost, they seem to have been conducted ethically with appropriate data and safety monitor controls. What is puzzling in retrospect is why, following publication of negative study results, the approach was not to re-examine the scientific premise but rather to make a design modification and then to proceed with another trial. Certainly, while unbiased peer review and oversight is important in checking this instinct going forward, it is equally important to establish that conflicts of interest, overt and subtle, among the key thought leaders driving the technology are acknowledged and accounted for.

Why wasn't the scientific community more overtly skeptical of the cardiac regenerative strategy? The answer might be that the larger community was in fact skeptical. In this regard, I am once again struck by the reproducibility literature whereby a "predictive market" of informed but not conflicted scientists can be engaged to predict the outcomes of reproducibility studies.[16] These markets perform extremely well and may be a low-cost approach to informing "go/no-go" decisions around new technologies. A predictive market was never engaged to evaluate the likelihood that an additional stem cell trial would be strikingly positive in the face of prior negative studies but it is certainly plausible that had it been, a healthy and pervasive skepticism might have manifested.

So where does this leave us? I would argue that there are lessons that can be learned: journals and journal editors inclined to publish high-profile/high-impact studies should demand statistical rigor, transparency around study registration, and prespecified analytical plans and whenever possible access to primary data. Reviewers and editors need to take their responsibilities to provide unbiased scientific oversight very seriously. When knowledge moves from discovery to clinical translation, there is even more need to evaluate overt and covert conflict, and whenever possible to divorce the financial implications of a project from its scientific integrity. Finally, it is intriguing to consider that studies that seem "too good to be true" when evaluated by a predictive market, might well be nonreproducible and that this needs to be rigorously tested. Scientific investigation is very hard to do and even in the best of circumstances, scientific conclusions are subject to reassessment in light of additional data. However, we live in an era when even the hint of impropriety can poison a dialogue and the risk is that bias can negate fact. In light of this, it is critical that we hold to fundamental principles of scientific rigor and integrity.

References

1. Kolata G. Harvard calls for retraction of dozens of studies by noted cardiac researcher. New York Times 2018:A13.
2. Kaiser J. Suspect science leads to pause in stem cell trial. Science 2018;362:513.
3. Epstein JA. A time to press reset and regenerate cardiac stem cell biology. JAMA Cardiol 2019;4:95–6.
4. Orlic D, Kajstura J, Chimenti S, Jakoniuk I, Anderson SM, Li B, et al. Bone marrow cells regenerate infarcted myocardium. Nature 2001;410:701–5.
5. Balsam LB, Wagers AJ, Christensen JL, Kofidis T, Weissman IL, Robbins RC. Haematopoietic stem cells adopt mature haematopoietic fates in ischaemic myocardium. Nature 2004;428:668–73.
6. Murry CE, Soonpaa MH, Reinecke H, Nakajima H, Nakajima HO, Rubart M, et al. Haematopoietic stem cells do not transdifferentiate into cardiac myocytes in myocardial infarcts. Nature 2004;428:664–8.
7. Glaser R, Lu MM, Narula N, Epstein JA. Smooth muscle cells, but not myocytes, of host origin in transplanted human hearts. Circulation 2002;106:17–9.
8. Clifford DM, Fisher SA, Brunskill SJ, et al. Stem cell treatment for acute myocardial infarction. Cochrane Database Syst Rev 2012;2:CD006536.
9. Maliken BD, Molkentin JD. Undeniable evidence that the adult mammalian heart lacks an endogenous regenerative stem cell. Circulation 2018;138:806–8.

10. Bergmann O, Zdunek S, Felker A, Salehpour M, Alkass K, Bernard S, et al. Dynamics of cell generation and turnover in the human heart. Cell 2015;161:1566–75.
11. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14:365–76.
12. Hauptman P, Chibnall JT. Unmeasured confounders and predictive models: what's your C-statistic? J Card Fail 2015;21:857–8.
13. Open Science Collaboration PSYCHOLOGY. Estimating the reproducibility of psychological science. Science 2015;349: aac4716.
14. Ioannidis JP, Munafò MR, Fusar-Poli P, Nosek BA, David SP. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci 2014;18:235–41.
15. Ebersole CR, Axt JR, Nosek BA. Scientists' reputations are based on getting it right, not being right. PLoS Biol 2016;14: e1002456.
16. Dreber A, Pfeiffer T, Almenberg J, Isaksson S, Wilson B, Chen Y, Nosek BA, Johannesson M. Using prediction markets to estimate the reproducibility of scientific research. PNAS 2015;112:15343–7.