Editorial

Replication in medical research: A modern day enigma

Sri Lanka Journal of Child Health, 2019; **48**(2): 97-99
DOI: http://dx.doi.org/10.4038/sljch.v48i2.8699
(Key words: Replication, Medical research)

Research is the cornerstone of progress in any form of science. The word research is defined as "the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions"[1]. More specifically, scientific research is aptly described as "the search for knowledge, or as any systematic investigation, with an open mind, to establish novel facts, solve new or existing problems, prove new ideas, or develop new theories, usually using a scientific method"[2]. The main resolve of basic research, in contradistinction to applied research, is discovering, interpreting and developing schemes for the improvement of human awareness on a variety of scientific staples of our world. Scrupulously performed scientific research is particularly important in healthcare because of the abiding implications that it has on diseases, their treatment and prevention. Over many a century, scholarly and scientific studies have provided a robust basis for the way we approach human diseases, unravel their complexities and decide on the best forms of their management. Without research, healthcare would stagnate and it would not be possible for us to move forward in a quest towards conquering many of the maladies that affect mankind.

Enter this enigma called replication. The very word 'replication', when used with reference to human investigation, is a quaint allusion to the duplication of a research study venture, generally with other diverse situations and different subjects, to assess if the primary results of the initial study can be confirmed and applied to other people and even other settings. Once a research study has been conducted, the investigators might be interested in evaluating whether the results hold true in different situations or for other diverse populations. In other cases, scientists may actually want to replicate the experiment to further demonstrate the results and confirm their original conclusions[3]. When replicating earlier endeavours, the investigators usually follow identical or very closely similar methodologies, but with a different set of participants in an altered time frame. If the new venture provides the same or similar results in follow-up experiments, it implies that the original results are less likely to be due to coincidence or to have occurred by chance. At the end of the day, replication is the mechanism through which credibility of scientific observations are confirmed. Put in another way, positive replication of scientific research is another base for progress.

Although this state of affairs is the one that is the expected norm and the pathway to progress, there seems to be a crunch situation in the closely-knit world of research. It is currently referred to as the 'Replication Crisis'. This refers to the well-documented fact that an unexpected quantity of published studies fail to replicate when the processes are repeated[4]. This may come as a shock to many but the truth is sometimes stranger than fiction.

Amongst quite a few efforts to assess reproducibility, two of them, the Reproducibility Project, a collaboration of 270 psychologists who attempted to replicate 100 psychology studies[5], and a 2018 report which examined studies printed in the most respected scholarly journals, Nature and Science between 2010 and 2015[6], are worthy of note. These efforts found that approximately two thirds of studies do replicate to some degree, but the robustness and strength of the findings are often weaker than originally claimed. It also means that a third of the replication endeavours fail to agree with the original index venture. As an extreme example, a study published in 2007 claimed that complicated maths problems needing careful thinking are easier to solve when presented in a fuzzy font[7]. However, 16 attempts to replicate that result were unsuccessful, indicating that the original claim was probably flawed[8]. When placed on a graph, the studies formed an impeccable bell curve centred on zero effect. In this instance, as one has frequently seen in similar circumstances, of all the attempts to replicate, the original index study had the smallest sample size as well as the most extreme result.

Most people erroneously accept that, due to the "$p<0.05$" cut-off for statistical significance, only 5% of discoveries will prove to be errors. In fact, John Ioannidis pointed out some faults in that contention[9]. He stated that false discoveries accounted for the major portion of the published research work. Apparently, recreations demonstrate that for a lot of study designs and the research scenarios, it is more likely for a study claim to be false than true. He also states that for many current scientific arenas, claimed study results may often be just the precise results of the prevalent technical

bias. Replication efforts do confirm that the false discovery rate is much higher than 5%. This should be of grave concern to all concerned.

Over 800 signatories of an article called for an end to overvalued claims and the summary dismissal of possibly crucial effects[10]. They categorically stated that we have to be clear as to what must be stopped. Their suggestions are that we should not conclude that there is 'no difference' or 'no association' just because a '*p*' value is larger than a cut-off level such as 0.05 or in the same vein, because a confidence interval includes zero. We should desist from concluding that two studies are at conflict because one had a statistically significant result and the other did not. These errors are a real waste of research efforts and they misinform the powers-that-be, who make policy decisions.

Well, is this good or bad for science? It is perhaps rather painful for several researchers whose work gets challenged. This is particularly relevant as the rate of failures of replication is perhaps way too high at the present stage of progress of the journal publication scenario. However, it must be recognised that these replication efforts are models of good operational science. They are concentrated examples of the use of rigorous scientific techniques, vigilant testing and surveillance in an admirable quest for reproducible and tangible outcomes. Rather than decline of science, it is endorsing the finest practices of the scientific method. However, 100% reproducibility cannot be assured for even the very best of studies. In that respect we do fall short of the 'Gold Standard'.

The replication crisis could be a blessing in disguise. Awareness of the replication crisis will hopefully promote better behaviour among scientists. Perhaps just a couple of decades ago, the cycle for publication was generally complete after an author or a set of authors managed to convince a couple of reviewers and an editor that the work was sound. However the research results would become a part of the published public record and a part of the scientific literature. Then it became open to further review by others but that was a very slow-moving process. However, in stark contrast, today's scientific world has raised the bar quite significantly for researchers. They should be well aware of the likelihood that their study or studies might be evaluated and carefully reviewed by a rather large number of opinionated and even narrow-minded observers on the World Wide Web. It may also happen with the involvement of high-profile groups like the lot in the Reproducibility Project.

Some journals now insist that scientists should make their data and computer codes available to those who ask for them. This makes it that much more likely for others to notice errors in their work. We, at the Sri Lanka Journal of Child Health, would publish original research only on the implied stipulation as indicated in our policies, that the authors would be prepared to submit the data for further evaluation whenever requested.

References

1. Definition of research. Available from https://en.oxforddictionaries.com/definition/research Accessed on 01-05-2019.

2. Research. Available from: http://mindbodyinstitutebeyond.com/research/ Accessed on 01-05-2019.

3. What is Replication? Available from: https://www.verywellmind.com/what-is-replication-2795802 Accessed on 02-05-2019.

4. The replication crisis is good for science. Available from: https://theconversation.com/thereplication-crisis-is-good-for-science-103736 Accessed 02-05-2019

5. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* 2015; **349**(6251): aac4716. https://doi.org/10.1126/science.aac4716 PMid: 26315443

6. Camerer CF, Dreber A, Holzmeister F, *et al*. Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nature Human Behaviour* 2018; **2**:637–44 https://doi.org/10.1038/s41562-018-0399-z

7. Alter AL, Oppenheimer DM, Epley NE, Rebecca N. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology* 2007; **136**(4):569-76. https://doi.org/10.1037/00963445.136.4.569 PMid: 17999571

8. Meyer AF, Shane B, Terence C, *et al*. Disfluent fonts don't help people solve math problems. *Journal of Experimental Psychology* 2015; **144**(2):e16-e30. https://doi.org/10.1037/xge0000049 PMid: 25844628

9. Ioannidis JPA. Why most published research findings are false. *PLOS Medicine* 2005; **2**(8): e124. https://doi.org/10.1371/journal.pmed.0020124
PMid: 16060722 PMCid: PMC1182327

10. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature* 2019; **567:** 305-7 https://doi.org/10.1038/d41586-01900857-9
PMid: 30894741

B J C Perera https://orcid.org/0000-0001-7789-8793
Joint Editor