



## Reproducibility in biomarker research and clinical development: a global challenge

“Implementation of an academic quality system will contribute to a better understanding of and trust in the quality and validity of research results...”

First draft submitted: 25 January 2017; Accepted for publication: 8 February 2017; Published online: 14 March 2017

**Keywords:** biomarkers • consortia • reproducibility

According to a recent survey conducted by the journal *Nature*, a large percentage of scientists agrees we live in times of irreproducibility of research results [1]. They believe that much of what is published just cannot be trusted. While the results of the survey may be biased toward respondents with interest in the area of reproducibility, a concern is recognizable. Goodman *et al.* discriminate between different aspects of reproducibility and dissect the term into ‘material reproducibility’ (provision of sufficient information to enable repetition of the procedures), ‘results reproducibility’ (obtaining the same results from an independent study; formerly termed ‘replicability’) and ‘inferential reproducibility’ (drawing the same conclusions from separate studies) [2]. The validity of data is threatened by many issues, among others by poor utility of public information, poor protocols and design, lack of standard analytical, clinical practices and knowledge, conflict of interest and other biases, as well as publication strategy.

### Publication bias

In 2007, Kyzas *et al.* found that almost all 1915 articles on cancer prognostic markers he looked at reported statistically significant positive associations [3]. Only about 1.4% of the articles were fully negative. The observed strong bias makes it very difficult to select and focus on promising biomarkers for follow-up studies. When Tzoulaki *et al.* analyzed PubMed entries for meta-analyses of cardio-

vascular biomarkers (excluding those that are part of the Framingham Study), they found an excess of ‘positive’ results [4]. Reasons for ‘selective analysis reporting bias’ are multiple hypothesis testing and certainly the degree of freedom that investigators have [5], for example, a study is conducted without hypotheses, or the hypothesis is changed during the study to find something ‘positive’ (hypothesis rescue bias). True ‘negative’ reports are rare, but important, such as the report on negative MRI biomarkers in the presence of strong tumor response to targeted cancer therapeutics by Boulton *et al.* [6]. Furthermore, incomplete reporting of studies makes it hard to estimate their value. McShane has called for improvement of quality, completeness and transparency of reporting of tumor biomarker studies to enhance their clinical utility [7]. Study preregistration is an interesting concept which has recently been developed and adopted by a number of journals ([8,9]). Under this strategy, the researcher submits the hypothesis and a research plan to the journal, where it is peer reviewed. Upon acceptance, the researcher has ease of mind that his manuscript will be accepted, no matter whether the outcome is positive or negative, as long as he/she adheres to the original project plan and passes a second review. In 2011, Andre *et al.* have proposed a registry for biomarker studies which would include details on the patient characteristics, biomarkers tested and results [10]. One may expand on this idea by implementing a



**Andreas Scherer**

Author for correspondence:  
Institute for Molecular Medicine  
Finland (FIMM), University of Helsinki,  
Helsinki, Finland  
and  
EATRIS ERIC, The European  
Infrastructure for Molecular Medicine,  
Amsterdam, The Netherlands  
Tel.: +358 50 448 7257  
[Andreas.Scherer@Helsinki.fi](mailto:Andreas.Scherer@Helsinki.fi)

preregistration database for clinical biomarker studies including the original hypothesis and its justification, as is done for therapeutic trials [11]. This could be done even on university/laboratory level, where biomarker studies (or any experiment for that matter) could be registered and locked. In fact, preregistration has been viewed as one of the ways out of the reproducibility crisis in the Nature survey mentioned earlier. In addition, the Nature survey respondents favor other ways to address the reproducibility crisis, for example, experiment standardization, better training and mentoring, and increasing the role of journals and funders in demanding thorough experiment design and result documentation and developing better reward systems for reproducibility studies. These are very important aspects and deserve thorough discussion and actions in the future. In my opinion, locking experiment design and hypothesis is important, and should be part of the implementation of academia quality system, similar to what already exists in industry. The Global Biological Standards Institute has linked the lack of such a system in academia to the reproducibility issue in academic research [12]. Implementation of an academic quality system will contribute to a better understanding of and trust in the quality and validity of research results coming from academia which industry largely builds on these days. By focusing on quality-ensured data from academia, industry should be able to reduce not only waste of time and money, but also limit unnecessary patient discomfort and suffering.

“...we cannot afford wasting time and resources on false positives or false negatives.”

### Standardization & harmonization efforts

Standardization efforts cannot be done in isolation if they are to reach public acceptance. Working in teams and in collaboration with other stakeholders, such as public–private partnerships or research consortia, provides a platform for exchange of ideas, analytical methods and better visibility to the stakeholders, than single laboratories can achieve. In the field of clinical biomarker development, method, material (e.g., cells and antibodies, but also software), and reporting standardization, identification of markers, validation, and application of biomarkers are so demanding processes that collaborations are vital [13]. For instance, development of biomarkers for monitoring the success of personalized interventions, a single, isolated organization will not be able to handle all operations. Through establishing such collaborations and consortia, it is possible to obtain expert advice in areas, that a single investigator would otherwise not have easily access to when working in isolation. With good management, a consortium

can address many aspects of reproducibility and generate new, relevant data and working hypotheses at the same time. Controlled application and comparison of methodologies, the development of standard operating procedures, preregistration and locking of hypotheses, and the potential to develop guidelines are only some of the benefits over working within an isolated laboratory. One example of such a collaboration is the US FDA initiated and -led community-wide effort for microarray quality control (MAQC, [14]), which started in 2005 to address the lack of guidelines and knowledge around the performance of microarray gene expression technologies. Teams from the USA, Europe, Asia and Australia came together to discuss and work on generating and analyzing high-throughput data. To address the question whether microarray gene expression data could be used as surrogate biomarkers of clinical end points, the analysis pipelines were designed and presented to a steering committee before the data analysis resumed, and locked for the rest of the study. Ideas and suggestions were shared among the analysis teams [15]. Meanwhile the MAQC has moved into the next-generation sequencing technology area and has assessed the reproducibility of technology platforms in the field of RNAseq [16] and recently in the genome sequencing, for example, for patient stratification (MAQCIV/SEQC2) [17]. Results and guidelines coming out of the MAQC/SEQC are valuable for the FDA and investigators alike, to evaluate the validity of data coming from the genomics community.

### Training

Nearly 90% of respondents of the Nature survey said they and the system would benefit from better training and mentorship [1]. I believe that all stakeholders involved in translational research, including investigators and funders, would benefit from objective mentorship and advice. Over recent years, the European Infrastructure for Translational Medicine (EATRIS) has taken on an additional role of a third-party advisory body to all stakeholders involved in translational research, including biomarker development. For instance, through the EATRIS expert panel of academic and clinical researchers and in response to the depth of the evaluation requirements of funding applications, EATRIS provides assessment of the clinical need addressed in a project proposal, identifies the relevant translational tools and facilities for optimal data generation, the clarity of the vision of the end product, the IP clarity and the regulatory pipeline [18]. These services provide tools to reduce research waste and toward improved use of public funding. To further address burning issues in the area of translational medicine, EATRIS with its more than 80 academic member

centers in 12 European states has realized the necessity of higher-grade collaborations across the globe. The 'Global Collaboration', involving the NIH National Center for Advancing Translational Science, USA; [19], Therapeutic Innovation Australia, AU; [20], MRC Technology (UK; [21]), EATRIS [22] and Center for Drug Research and Development; [23], is a novel endeavor to jointly utilize complementary expertise from all areas in translational research, with the goal of expediting the development of quality therapeutic interventions, and focus on exploring the criteria for reproducibility of key technologies in biomedical research [24].

## Conclusion

At large, the current system of science is determined by financial aspects and a pressure to publish on scientists and journals alike. Funders are often overwhelmed by the sheer number of applications and may tend toward accepting overclaiming proposals.

Literally, we cannot afford wasting time and resources on false positives or false negatives. Understanding the relationship between method and interpretation reproducibility with all its implications in statistics and clinical knowledge will be essential. We

need to improve training in the necessity of ensuring data and reporting quality for successful translational and biomarker research, demand and create formal quality systems for academic research and encourage active participation in international collaborations. Adherence to reporting guidelines must be required, not only recommended. Through activities of organizations, such as EATRIS, funders can get support in shortlisting applications, and investigators can strongly improve the quality of their applications by getting access to external, objective expertise in all areas of translational research.

It is time for changes and we need to act in order to save cost, time and, most importantly, human lives.

## Financial & competing interests disclosure

The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

## References

- Baker M. 1,500 scientists lift the lid on reproducibility. *Nature* 533(7604), 452–454 (2016).
- Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? *Sci. Transl. Med.* 8(341), 341ps312 (2016).
- Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Almost all articles on cancer prognostic markers report statistically significant results. *Eur. J. Cancer* 43(17), 2559–2579 (2007).
- Tzoulaki I, Siontis KC, Evangelou E, Ioannidis JP. Bias in associations of emerging biomarkers with cardiovascular disease. *JAMA Intern. Med.* 173(8), 664–671 (2013).
- Wicherts JM, Veldkamp CL, Augusteijn HE, Bakker M, van Aert RC, van Assen MA. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-hacking. *Front. Psychol.* 7, 1832 (2016).
- Boult JK, Jamin Y, Jacobs V *et al.* False-negative MRI biomarkers of tumour response to targeted cancer therapeutics. *Br. J. Cancer* 106(12), 1960–1966 (2012).
- Mcshane LM, Hayes DF. Publication of tumor marker research results: the necessity for complete and transparent reporting. *J. Clin. Oncol.* 30(34), 4223–4232 (2012).
- E Pain. Register your study as a new publication option. *Science*. [www.sciencemag.org/](http://www.sciencemag.org/)
- Munafò M. Open science and research reproducibility. *Ecancermedicalscience* 10, ed56 (2016).
- Andre F, Mcshane LM, Michiels S *et al.* Biomarker studies: a call for a comprehensive biomarker study registry. *Nat. Rev. Clin. Oncol.* 8(3), 171–176 (2011).
- ClinicalTrials.gov. [www.clinicaltrials.gov](http://www.clinicaltrials.gov)
- Global Biological Standards Institute. [www.gbbsi.org/publication/the-case-for-biological-standards/](http://www.gbbsi.org/publication/the-case-for-biological-standards/)
- Asadullah K, Busch A, Gottwald M, Reinke P, Landeck L. Industry-academia collaborations for biomarkers. *Nat. Rev. Drug. Discov.* 14(12), 805–806 (2015).
- US Food & Drug Administration. [www.fda.gov/ScienceResearch/](http://www.fda.gov/ScienceResearch/)
- Shi L, Campbell G, Jones WD *et al.* The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat. Biotechnol.* 28(8), 827–838 (2010).
- Xu J, Gong B, Wu L, Thakkar S, Hong H, Tong W. Comprehensive assessments of RNA-seq by the SEQC consortium: FDA-led efforts advance precision medicine. *Pharmaceutics* 8(1), pii:E8 (2016).
- Xu J, Thakkar S, Gong B, Tong W. The FDA's experience with emerging genomics technologies-past, present, and future. *AAPS J.* 18(4), 814–818 (2016).
- van Dongen GA, Ussi AE, de Man FH, Migliaccio G. EATRIS, a European initiative to boost translational biomedical research. *Am. J. Nucl. Med. Mol. Imaging* 3(2), 166–174 (2013).
- National Center for Advancing Translational Sciences. <https://ncats.nih.gov>
- Therapeutic Innovation Australia. [www.therapeuticinnovation.com.au](http://www.therapeuticinnovation.com.au)
- MRC Technology. [www.mrc technology.org](http://www.mrc technology.org)

- 22 EATRIS.  
[www.EATRIS.eu](http://www.EATRIS.eu)
- 23 The Centre for Drug Research and Development.  
[www.CDRD.ca](http://www.CDRD.ca)
- 24 Gilliland CT, Zuk D, Kocis P *et al.* Putting translational science on to a global stage. *Nat. Rev. Dru. Discov.* 15(4), 217–218 (2016).