iScience

CellPress
OPEN ACCESS

Perspective

Assessing reproducibility of the core findings in cancer research

Hassan Fazilaty[1,2,*]

INTRODUCTION

Cancer is a leading cause of death worldwide (Siegel et al., 2022). Consequently, basic and preclinical cancer biology are intensively researched. This is essential to understand and then find treatments for different cancer types; however, the reproducibility of the new findings needs to be carefully measured and openly communicated. The burst of discoveries in cancer research during few past decades has challenged this notion. On one hand, the complexity of experimental approaches as well as biological systems and on the other hand, competition for publication has caused tremendous difficulties to follow the reliability of new discoveries. There will be no benefit for patients or the public, if the studies are not reproducible. This is crucially important as the majority of the studies use public funding. To be able to rely on results from cancer studies for potential new treatments, the scientific community needs to find ways to measure reproducibility in a reliable manner.

More than 90% of cancer-related mortality is due to metastasis, which is a multistep and complex process (Hanahan and Weinberg, 2011). Metastasis is the spread of tumor cells from their primary site to secondary organs. During this journey, cancer cells encounter a variety of interactions with other cells and plenty of environmental cues. The different cell interactions include direct or indirect crosstalk with immune, endothelial, fibroblast, and other resident cells in every given tissue. There are different types of cells in every tumor making a heterogeneous makeup that is different from patient to patient. Tumor cells invade the surrounding tissue, intravasate into the circulation, some of them survive and reach secondary organs, extravasate into the recipient tissue, some survive, proliferate, and make new tumors (Valastyan and Weinberg, 2011). Cancer researchers have been vastly dedicated to understand this complexity by building models and breaking this complex problem into smaller/more understandable problems to be able to solve it and make efficient treatments.

Cell lines are among the simplest models that are vastly used as *in vitro* systems to study cancer biology and to test drugs. Cell lines are generally easy, quick, and cheap to work with. They are used to dissect molecular mechanisms by manipulating genes and signaling pathways in 2D or 3D culture systems. In addition, cell lines are widely used to test a variety of therapeutic options including chemical and biological drugs. Many so called "*in vivo*" systems also rely on using cell lines, for instance, injecting manipulated cells into mice and assessing the tumor growth and therapy response. Although closer to real physiological/pathological states, these models also suffer from the inherent limitations of cell lines as they are originated *in vitro* systems. A true *in vivo* system would require the disease (e.g. tumor) to appear spontaneously in an animal model (e.g. mouse) and progress in a similar manner compared to human disease. The issue is that most cell lines cannot recapitulate the complexity and heterogeneity of the original tumor. The different environmental cues (culture system) and cellular interactions *in vitro* compared to *in vivo* result in drastic changes in the makeup of cells extracted from a tumor. In addition, in order to keep primary tumor cells extracted from a human tumor (or mouse tissue) in culture for long time, they need to be transformed (e.g. by oncogenic viral genes). All these issues add up and make cell lines artificial systems. This does not mean that cell lines have no benefit, on the contrary, they have helped us understand many molecular mechanisms and features of cancer cells. However, to reach a comprehensive understanding (e.g. whether a therapy works for a cancer type), using a small number of cell lines *in vitro*, or even injecting them into mice is most likely not enough.

In an attempt to measure reproducibility in cancer research, a project was launched to openly investigate a set of studies that had been published in high-profile journals (Errington et al., 2014). The majority of experimental designs that were chosen for this project relied on cell lines *in vitro*, and in some cases injecting cell lines in mice. The final report of the project consists of data from 50 replication experiments covering 23 original studies. Comparing the replication studies to those of the original papers, they found that the

[1]Department of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[2]Center for Reproducible Science (CRS), University of Zurich, Zurich Switzerland

*Correspondence: hassan.fazilaty@uzh.ch

https://doi.org/10.1016/j.isci.2022.105125

replications were 85% weaker in median effect size (Errington et al., 2021c). The weaker evidence, which was observed for both *in vitro* and *in vivo* experiments calls for additional attention to delineate the challenges of replicability and the necessity to improve transparency and rigor in research practices. This pioneering attempt to carefully measure reproducibility revealed the challenges of designing and conducting replication studies (Errington et al., 2021a). The fact that this project was unable to perform and reproduce a major part of the planned replication studies might seem extremely worrying (Mullard, 2021). However, a deeper look at the attempt is needed to realize the source of the challenges, reinforcing the need to find better ways to measure reproducibility. In this short piece, I will point out the main issues of the replication studies and suggest a new approach to assess reproducibility of major biological studies including cancer research.

CHALLENGES AND ISSUES WITH REPRODUCIBILITY OF EXPERIMENTS IN CANCER RESEARCH

In the most recent, and final, report from the Reproducibility Project: Cancer Biology, Errington and colleagues conducted experiments for 11 unfinished registered reports (Errington et al., 2021b). Four papers were excluded. The replication of the remaining studies was incomplete due to technical or methodological challenges that the authors did not anticipate. The main reasons were methodological challenges, the complexity of the approaches, and limitations in funding. In principle, the replication attempts lacked required expertise and resources to truly replicate the original studies. The core experiments that lead the authors of the original studies to their conclusions were included, partially, in the registered reports for replication but, disappointingly, in most cases, were excluded from the experimental work in the final report.

Excluded parts were *in vivo* experiments including using mouse models, immune staining using antibodies, or more complex *in vitro* experiments such as 3D cell culture (Errington et al., 2021b). A telling example is the attempted replication of the study by Ricci-Vitiani and colleagues. The replication study only used a subset of the cell lines tested in the original publication. Crucially, not tested were GNS (glioblastoma neurospheres) cells, that in the original study responded substantially to treatment (Ricci-Vitiani et al., 2010). Unfortunately, *in vivo* experiments were also not replicated because the authors could not efficiently generate the required cell line to be injected into mice (Errington et al., 2021b). The absence of the *in vivo* experiments undermines the value of the replication. Indeed, the value of the few experiments that were (partially) successfully conducted in the final replication study is questionable.

As well as excluding key experiment, in the final reproducibility study report, different reagents or tools were often used. The substituted experimental approaches are problematic alternatives. For instance, to replicate the results from Ricci-Vitiani and colleagues, the Reproducibility Project used method of choice was quantitative real-time PCR in the replication study while in the original paper flow cytometry was used (Ricci-Vitiani et al., 2010). The method changes means that transcript levels were analyzed instead of protein. Another example is the attempt to repeat the experiments from Heidorn and colleagues (Heidorn et al., 2010). The authors used SB590885 compound (BRAF inhibitor) in the replication study while another inhibitor 885A was used in the original study. Notably, the authors mention that SB590885 is a close analog of 885A; however, the only non-significant finding is observed after using this inhibitor. It needs to be considered that the change contributed in the difference. In another example, to replicate the findings of Johannessen and colleagues (Johannessen et al., 2010), the experimental system of choice contained a different cell line in the replication study: HT-29 cells instead of OUMS-23 colon cancer cells. These deviations in experimental settings and basing the replication studies solely on *in vitro* systems can exert strong confounding effects on the results and therefore render the comparison unreliable.

In the majority of cases, the registered reports were designed following the original reports and the deviations occurred during the replication studies. However, the design of the replication studies is another crucial point that needs to be taken into account. For instance, the use of 885-A instead of SB590885 is already included in the registered report (Bhargava et al., 2016). The question is why changes from the original protocol may have occurred. Were the changes missed during the review process, or were they considered negligible? On the hand, the deviations from the original studies might be construed as improvements. These issues could create additional layers of misunderstanding and misconception and might confound finding the source of irreproducibility. The main reason for these types of aberrations could be the lack of understanding of the core principles and critical features of the original research.

The difficulty of fully replicating the selected studies is taken by the authors as evidence that reproducing and/or replicating the preclinical cancer biology experiments is challenging (Errington et al., 2021b). However, it is a challenge that must be met if concerns about reproducibility are to be resolved. Replicability studies may be unreliable or uninformative if the most important part of a study is not proficiently replicated in well-controlled settings. On the other hand, in the majority of the cases, reliance on a single *in vitro* model is not sufficient to draw a conclusion about a biological phenomenon as well as measuring the effect of a treatment. This applies both to original research and replication efforts.

Assessing the situation form a different viewpoint, it needs to be considered that maybe not all the issues discussed about the replication studies could be anticipated in advance. It can be easy to spot the factors in the replication studies after they are done, which may explain why the results are different from the original studies. This can particularly occur when the authors of the original studies review the replication studies. While there is understandable motivations to believe that the original study was conducted flawlessly, it needs to be taken into account that the found problems in the replication studies are mainly speculative. The speculations can be valuable and lead to new research. However, neither the replication studies nor the original authors' responses can provide definitive evidence about the original studies. Nonetheless, these communications point to the fact that additional research is needed to stablish reliability of the findings.

FUTURE DIRECTIONS AND CONCLUSION

Reproducibility and replicability can be best measured if several model systems, particularly true *in vivo* models, are used. *In vitro* systems, particularly immortalized cell lines, need to be used with caution. Different passage numbers, or even minor changes in culture conditions, such as reagents in the medium or the density of the cells, can affect the experimental outcome. In cancer cell lines, these effects may be particularly bigger due to their high mutational burden and chromosomal defects. There are other examples of replication studies that were not able to reproduce the data from original publications while relying on cell lines (Yan et al., 2019). Although cell lines are excellent tools to explore molecular mechanisms or to perform large-scale screens, the sole use of them may not yield sufficient biological and/or physiological weight to draw conclusions.

A different setting might help better assess reproducibility of complex cancer projects. The limitations of the reproducibility project, particularly the final report, include i) excluding the most crucial experiments, ii) relying on a limited *in vitro* systems, and iii) failing to use the exact same conditions. These issues mainly come from lack of expertise and resources. Many *in vivo* and *in vitro* models require a serious investment of time to master—in the order of months and sometimes years. A collaborative effort between established expert labs to replicate parts of published experiments would be a better solution. For fields such as cancer, research labs with the requisite know-how exist. To assess reproducibility in an efficient way, such labs should be invited to perform the experiments.

One big issue in reproducibility comes from incomplete communication of protocols and data in original studies. In recent years, there are great improvements in this regard; however, this has been one of the major problems that the Reproducibility Project encountered since the beginning (Rodgers and Collings, 2021). The incomplete information at the time of publication can sometimes be difficult to spot if the editors/reviewers are not deeply familiar with procedures. Acquiring the collaborative strategy in which expert labs are in charge of performing the experiment can also overcome this issue in large part, as the experiments subjected to replication may be routinely done in this situation.

Having said that, I believe the best measure of reproducibility and replicability can come from the original studies. Authors of basic and preclinical cancer studies can best illustrate this notion by using several experimental systems, particularly *in vivo* validations as well as confirmations using human-derived systems. This is already the case for many publications and may be extended by better information and more rigorous and open peer review. Counterarguments of the authors from original publications also point in this direction that many details in performing experiments are acquired after a long time of focus and trials and some of these experiences might be extremely challenging to be replicated entirely and exactly (Mullard, 2021). Additionally, pre-registered reports can bring more visibility and robustness to preclinical studies and avoid unnecessary competition. In conclusion, to reach better treatment options for complex diseases like cancer, we need reliable, replicable, and reproducible science and this can be achieved by rigorous experimentation and expert examination of replicability.

ACKNOWLEDGMENTS

I am grateful to all members of the Basler lab for stimulating discussions, Konrad Basler for immense support, and George Hausmann for valuable comments on the manuscript. I also acknowledge Leonhard Held for valuable comments and support, as well as members and fellows of the Center for Reproducible Science of the University of Zürich for their input. My work has been supported by the Forschungskredit of the University of Zürich grant FK-21-119, and a grant of the University and Medical Faculty of Zurich and the Comprehensive Cancer Center Zurich.

DECLARATION OF INTERESTS

The author declares no competing interest.

REFERENCES

Bhargava, A. (2016). Registered report: Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. eLife. https://doi.org/10.7554/eLife.11999.

Errington, T.M., Iorns, E., Gunn, W., Tan, F.E., Lomax, J., and Nosek, B.A. (2014). An open investigation of the reproducibility of cancer biology research. Elife *3*, e04333. https://doi.org/10.7554/eLife.04333.

Errington, T.M., Denis, A., Perfito, N., Iorns, E., and Nosek, B.A. (2021a). Challenges for assessing replicability in preclinical cancer biology. Elife *10*, e67995. https://doi.org/10.7554/eLife.67995.

Errington, T., M. (2021). Investigating the replicability of preclinical cancer biology. eLife. https://doi.org/10.7554/eLife.71601.

Errington, T.M., Denis, A., Allison, A.B., Araiza, R., Aza-Blanc, P., Bower, L.R., Campos, J., Chu, H., Denson, S., Donham, C., et al. (2021b). Experiments from unfinished registered reports in the reproducibility project: cancer biology. Elife *10*, e73430. https://doi.org/10.7554/eLife.73430.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674. https://doi.org/10.1016/j.cell.2011.02.013.

Heidorn, S.J., Milagre, C., Whittaker, S., Nourry, A., Niculescu-Duvas, I., Dhomen, N., Hussain, J., Reis-Filho, J.S., Springer, C.J., Pritchard, C., and Marais, R. (2010). Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. Cell *140*, 209–221. https://doi.org/10.1016/j.cell.2009.12.040.

Johannessen, C.M., Boehm, J.S., Kim, S.Y., Thomas, S.R., Wardwell, L., Johnson, L.A., Emery, C.M., Stransky, N., Cogdill, A.P., Barretina, J., et al. (2010). COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. Nature *468*, 968–972. https://doi.org/10.1038/nature09627.

Mullard, A. (2021). Half of top cancer studies fail high-profile reproducibility effort. Nature *600*, 368–369. https://doi.org/10.1038/d41586-021-03691-0.

Ricci-Vitiani, L., Pallini, R., Biffoni, M., Todaro, M., Invernici, G., Cenci, T., Maira, G., Parati, E.A., Stassi, G., Larocca, L.M., and De Maria, R. (2010). Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells. Nature *468*, 824–828. https://doi.org/10.1038/nature09557.

Rodgers, P., and Collings, A. (2021). What have we learned? Elife *10*, e75830. https://doi.org/10.7554/eLife.75830.

Siegel, R.L., Miller, K.D., Fuchs, H.E., and Jemal, A. (2022). Cancer statistics, 2022. CA. Cancer J. Clin. *72*, 7–33. https://doi.org/10.3322/caac.21708.

Valastyan, S., and Weinberg, R.A. (2011). Tumor metastasis: molecular insights and evolving paradigms. Cell *147*, 275–292. https://doi.org/10.1016/j.cell.2011.09.024.

Yan, X., Tang, B., Chen, B., Shan, Y., and Yang, H.; Reproducibility Project: Cancer Biology (2019). Replication Study: the microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Elife *8*, e43511. https://doi.org/10.7554/eLife.43511.