Towards best practices in research

Role of academic core facilities

Leonardo Restivo[1,*,†] [iD], Björn Gerlach[2,**,†] [iD], Michael Tsoory[3] [iD], Lior Bikovski[4,5], Sylvia Badurek[6] [iD], Claudia Pitzer[7] [iD], Isabelle C Kos-Braun[7] [iD], Anne-Laure MJ Mausset-Bonnefont[8] [iD], Jonathan Ward[9], Michael Schunn[10], Lucas PJJ Noldus[11,12] [iD], Anton Bespalov[2] & Vootele Voikar[13,***] [iD]

During the past decade, the scientific community and outside observers have noted a concerning lack of rigor and transparency in preclinical research that led to talk of a "reproducibility crisis" in the life sciences (Baker, 2016; Bespalov & Steckler, 2018; Heddleston *et al*, 2021). Various measures have been proposed to address the problem: from better training of scientists to more oversight to expanded publishing practices such as preregistration of studies. The recently published EQIPD (Enhancing Quality in Preclinical Data) System is, to date, the largest initiative that aims to establish a systematic approach for increasing the robustness and reliability of biomedical research (Bespalov *et al*, 2021). However, promoting a cultural change in research practices warrants a broad adoption of the Quality System and its underlying philosophy. It is here that academic Core Facilities (CF), research service providers at universities and research institutions, can make a difference.

It is fair to assume that a significant fraction of published data originated from experiments that were designed, run, or analyzed in CFs. These academic services play an important role in the research ecosystem by offering access to cutting-edge equipment and by developing and testing novel techniques and methods that impact research in the academic and private sectors alike (Bikovski *et al*, 2020). Equipment and infrastructure are not the only value: CFs employ competent personnel with profound knowledge and practical experience of the specific field of interest: animal behavior, imaging, crystallography, genomics, and so on. Thus, CFs are optimally positioned to address concerns about the quality and robustness of preclinical research.

The role of core facilities

Core Facilities create several benefits to the scientific community above providing access to equipment and expertise. Shared access to research infrastructure encourages collaborations within institutions and beyond (Gould, 2015). The CF and its personnel are often a central knowledge hub where different research groups converge and interact. Many CFs store large data volumes, including negative or confirming data, which would benefit the

scientific community at large if it were accessible (Bespalov *et al*, 2019). Finally, CFs and their personnel are actively teaching and training users specialized knowledge from sample preparation to experiment design to data analysis. These training activities, whether individually or in groups, are a good opportunity for educating future generations of scientists about crucial concepts of research rigor, responsible conduct of research, and reproducibility.

....

" . . . CFs are optimally positioned to address concerns about the quality and robustness of preclinical research."

....

Indeed, CFs often have difficulties to ensure that its users follow best research practices (Sherman, 2018; Knudtson *et al*, 2019; Kos-Braun *et al*, 2020). There are several reasons: insufficient training in research ethics or lack of knowledge of best practices (Festing, 2013); users arguing that their experimental procedures are well established in their own research units and that

1 Neuro-BAU, Department of Fundamental Neuroscience, Faculty of Biology & Medicine, University of Lausanne, Lausanne, Switzerland
2 PAASP GmbH, Heidelberg, Germany
3 Behavioral and Physiological Phenotyping Unit, Department of Veterinary Resources, Weizmann Institute of Science, Rehovot, Israel
4 The Myers Neuro-Behavioral Core Facility, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
5 School of Behavioral Sciences, Netanya Academic College, Netanya, Israel
6 Preclinical Phenotyping Facility at Vienna BioCenter Core Facilities (VBCF), member of the Vienna BioCenter (VBC), Vienna, Austria
7 Interdisciplinary Neurobehavioral Core, University of Heidelberg, Heidelberg, Germany
8 IRMB, University of Montpellier, INSERM, Montpellier, France
9 Genome Biologics, Frankfurt am Main, Germany
10 Institute of Science and Technology Austria (IST Austria), Klosterneuburg, Austria
11 Noldus Information Technology BV, Wageningen, The Netherlands
12 Department of Biophysics, Radboud University, Nijmegen, The Netherlands
13 Neuroscience Center and Laboratory Animal Center, Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland
   *Corresponding author. E-mail: leonardo.restivo@unil.ch
   **Corresponding author. E-mail: bjoern.gerlach@paasp.net
   ***Corresponding author. E-mail: vootele.voikar@helsinki.fi
   †These authors contributed equally to this work

any change might render the result incompatible with previous studies (Lazic, 2016); and a lack of authority on the side of CFs to impose best practices (Kos-Braun *et al*, 2020). Of note, most CFs do provide training and consulting on best practices and monitor the experiments and the quality of data generated. However, enforcing best practices in all cases is often an unrealistic goal owing to the large volume of data generated by many different users (Kos-Braun *et al*, 2020) and the limited time and resources to explain users best practices and standards.

A memorandum of understanding

In the context of the EQIPD project, a working group (EQIPD-CF-WG) of 13 researchers (3 *in vitro*, 1 IT, 9 *in vivo*) drafted a Memorandum of Understanding (MoU) with the aim of applying the EQIPD core requirements (Bespalov *et al*, 2021) to CFs and their users. The central tenet of the MoU revolves around who is "responsible" to ensure and follow best research practices and what information users need to help them improve the quality of their experiments and the data thus generated.

The MoU, which is available as an Appendix and on the Open Science Framework webpage (https://osf.io/uk8wf), is structured in 3 sections (*Background*, *Recommendations,* and *Glossary*). The *Background* section describes the intention to improve the robustness, reliability, traceability, and integrity of data produced in academic CFs. It introduces two types of services offered by CFs—*Regular Service* or *EQIPD service*—and explains how the *Recommendations* to improve data quality could be adopted by CF users in their research.

The *Regular Service* addresses potential bias by encouraging users to incorporate the *Recommendations* in their workflow. It clearly defines that the entire responsibility and accountability for the data generated by the CF is with the user. The *EQIPD Service* assumes that the CF takes on a broader responsibility by implementing relevant recommendations and monitoring data quality, based on the EQIPD Quality System. CF personnel will provide practical advice throughout the research project, relating to experimental design, data generation, analysis, and documentation and will perform spot checks on the quality practices applied by the user. Following the completion of the project, users get formal approval that the

research was "EQIPD compliant" in the form of a seal of quality or a "badge".

The *Recommendations* section applies to both Regular and EQIPD services, and it offers practical guidelines, defines the responsibilities of, and expectations from, the user. These guidelines are an "agreement" between CFs and users, and they include topics such as training, experimental record, rigor in study design, data analysis, storage and traceability, and review and reporting (see Appendix or https://osf.io/uk8wf). Applying these *Recommendations* will ensure more robust research. Opting for the *EQIPD Service* would result in reliable reporting of research output—scientific and supporting evidence—in different types of publications: peer-reviewed articles, presentations, reports. The short *Glossary* at the end of the MoU aims to provide a common understanding on some important aspects of the research process.

CFs response to the memorandum

As the MoU was developed by a small group of CF heads, we sought feedback from the broader community and sent it for review along with a short online survey to 1,053 CFs. The survey asked CF leaders questions related to two main issues: Do they think that the MoU could improve research data? And what do they think may motivate users to adhere to best research practices guidelines? A total of 172 CFs replied with comments and feedback (Fig 1). The participants had diverse backgrounds and overall provided different types of service (full-, hybrid, and self-service) (Fig 1A and B). We found that a considerable proportion of respondents reported difficulties to motivate users to follow best practices (Fig 1C and D); only a minority (16.4%) strongly disagreed with the statement that it is challenging to motivate users to apply best practices. The "comments" indicated though that some CFs already apply drastic measures to the extent of denying "*access to facility itself (we do not allow external to use instrumentation if they don't follow trainings provided by us and implemented SOPs)* [...]" while others refer users to effective programs for "*training young scientists (PhD Students) and make them aware of the notion of 'best practice*".

Of course, under optimal conditions when sufficient resources are available, CFs will have little difficulties in motivating users to

adopt best practices. However, these conditions seldom occur; therefore, a different model must be envisaged to improve research data integrity. Although some respondents exhibited a very pessimistic view and labeled researchers who do not follow best practices as "*irredeemable*", respondents also acknowledged that there are "*few who need motivation*" and "*the vast majority of scientists care enormously about good science*".
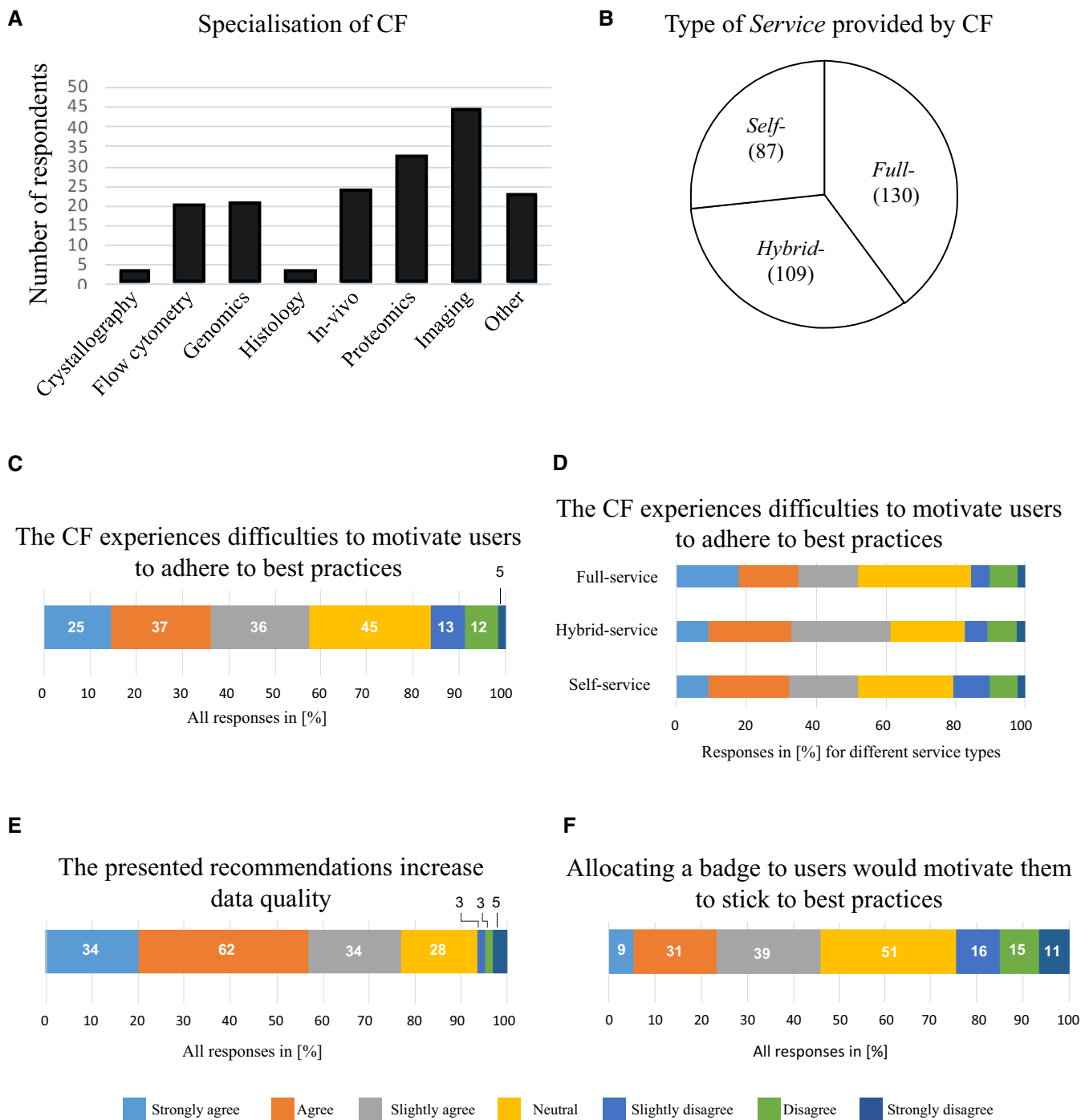
The feedback from the CF heads on the impact the MoU may have on the quality of preclinical research data was overall very positive (Fig 1E). Importantly, the attitude appears to be more positive in the case of hybrid service, where the user usually is supervised by the CF personnel. Additionally, the survey assessed whether CFs believe it is possible to motivate the users to follow best practices, as described in the MoU, by granting a badge (Fig 1F).

> " *. . . wide recognition of such a badge may provide the support and encouragement to implement best research practices.* "

A major concern is that the badge may become effective only if it is recognized on a larger scale, for example, by journal publishers and funding agencies—if "*journals force them to*" or "*Journal and funders asking for proof of adherence to good practices* [...]". Indeed, wide recognition of such a badge may provide the support and encouragement to implement best research practices. In addition, survey respondents suggested different ways to motivate users to adhere to best practices ("*Constant reminding! Along with easy to access guides detailing best practice*", "*The members of the facility should follow the best practices which would influence the users directly or indirectly*".). Alternatively, more robust supervision of early-stage career researchers was proposed ("*more supervision and guidance by their scientific supervisors*").

Stories from the working group

The EQIPD-CF-WG agrees with the claim by Kos-Braun and colleagues that a lack of responsibility, that is, who is responsible for the quality of the data, is often the cause of many problems that afflict the overall robustness of data collected in a CF (Kos-

A Specialisation of CF

B Type of *Service* provided by CF

C The CF experiences difficulties to motivate users to adhere to best practices

All responses in [%]

D The CF experiences difficulties to motivate users to adhere to best practices

Responses in [%] for different service types

E The presented recommendations increase data quality

All responses in [%]

F Allocating a badge to users would motivate them to stick to best practices

All responses in [%]

Strongly agree ■ Agree ■ Slightly agree ■ Neutral ■ Slightly disagree ■ Disagree ■ Strongly disagree

Figure 1. Responses to the survey from 172 Core Facilities (CF).

(A) Specialization of CF, (B) service provided by CF, (C) responses to the question whether CFs have troubles motivating users, (D) questions as in (C) but responses according to the type of service provided; (E) responses to the question whether the recommendations would increase data quality; (F) responses to the question whether a badge would motivate users to apply best practices, the results are shown for all CF and absolute numbers are given in brackets. Absolute numbers are reported in brackets (panel B) or inside the bars (panels C, E, and F).

Braun *et al*, 2020). Whenever the CF offers full service, the responsibility for data collection and management is completely attributed to the CF. However, when the CF provides self- or hybrid service, it is often not clear who is responsible. We found that the lack of transparency in respect to responsibility often negatively influences the overall robustness of the data on different levels from traceability to rigor in study design. Here, we provide a few "case studies" to illustrate the value of the recommendations to avoid bias in experimental design, encourage the application of rigorous quality standards, and improve the

traceability of research data supported by academic CFs.

> ❝ *We found that the lack of transparency in respect to responsibility often negatively influences the overall robustness of the data on different levels from traceability to rigor in study design.* ❞

Research projects can often span many years from collecting the first datasets to writing up a PhD thesis or journal article. This requires a system to guarantee the stability, traceability, and uncorrupted storage of data either at the CF or at the research unit. However, research units in academic environments are dynamic entities with a consistent turnover of personnel of students and postdocs that can even relocate to a different institution. Moreover, research projects can require several rounds of revision before the final publication. It is therefore critical to clarify who is responsible for both data collection and long-term data storage to prevent any adverse outcomes.

By way of example, a CF collaborated with a research unit that was later spun off into a company. A few years later, a prospective buyer wanted to make sure that no other party could claim the intellectual property. To ensure this, the original raw data had to be unequivocally owned by either the CF or the research unit. Unfortunately, neither was in possession of the raw data since both parties assumed that it was the responsibility of the other to store it. Consequently, the acquisition was halted.

To prevent such misunderstandings, the MoU requires that the CF must clarify early on who is responsible for the overall data quality including traceability. The MoU also provides suggestions on best practices to make data traceable and identifiable across time and personnel turnover and introduces the concept of experimental record to collect data in a structured and traceable format (https://eqipd-toolbox.paasp.net/wiki/Experimental_Record).

Rigor in study design includes different practices aimed at improving the confidence in the data and the conclusions drawn from it (Bespalov & Steckler, 2018; Turkiewicz *et al*, 2018; Heddleston *et al*, 2021). The MoU provides a 3-level structured list of important

prescriptions: those that must be followed—such as hypothesis, description of sample size, inclusion/exclusion criteria; recommendations that should be followed, for instance the choice of experimental methods; and advice, for instance, protocol preregistration.

Yet, the following two scenarios are unfortunately still too common. A researcher finds an interesting result in an exploratory experiment and decides to keep increasing the sample size to reach statistical significance. Albeit wrong according to statistical frequentist approaches, this is an all-too-common practice that was reported by many CF heads who took part in the working group. To address this, the MoU supports constructive communication between the CF and the users by reminding users that suitable sample size to assess the effect of exploratory research must be computed *a priori*; and via the EQIPD toolbox (https://eqipd-toolbox.paasp.net/wiki/Toolbox) to support the researcher in the development of robust experimental designs.

> ❝ *The feedback to these guidelines was to a major extent positive and our peers [...] agreed on the potential effectiveness of the MoU, if followed by the users, to achieve its aim.* ❞

In another case, a researcher asks for the removal of outliers to reach an "almost significant difference". In addition, it is not infrequent that researchers adopt one-tailed tests to compute P-values, leading to an inflation of the false-positive rate. However, innocuous these practices may sound, they compound to an elevated degree of inaccuracy with catastrophic consequences for the claims and conclusions drawn from the data. The root of the problem may be attributed to both misinterpretation and the lack of knowledge of the mechanics that govern the frequentist approach to statistics (Nuzzo, 2014; Greenland *et al*, 2016) which is a key contributor to irreproducibility of biomedical data (Turkiewicz *et al*, 2018). Indeed, the recommendations cannot possibly cover all scenarios, but they do highlight in the *Rigor in Study Design* section the main points that we found at the root of the many poor practices we observed in CFs.

Where to go next?

We developed a Memorandum of Understanding between CFs and their users with the aim to improve the overall data quality of preclinical research. The recommendations range from experiment logging, rigor in study design to data analysis, review, and reporting. The feedback to these guidelines was to a major extent positive, and our peers with different backgrounds agreed on the potential effectiveness of the MoU, if followed by the users, to achieve its aim. The MoU offers users a choice between the *Regular Service* and the *EQIPD Service*. Both choices clearly state at the outset of the collaboration who is responsible for assessing the quality of the data. While the locus of responsibility for overseeing the quality of the data should not affect data integrity, we realize that having the CF playing a more prominent role in the research process and overseeing data quality may offer further protection from bias. In addition, CFs are ideally suited to promote best practices since they are usually independent and impartial to the pressure to publish original findings and compelling stories.

> ❝ *... CFs are ideally suited to promote best practices since they are usually independent and impartial to the pressure to publish original findings and compelling stories.* ❞

To this end, the CF may independently implement the EQIPD Quality System (Bespalov *et al*, 2021) and offer the *EQIPD Service*. It allows the CF to identify the best solutions to implement the recommendations; perform spot checks on data integrity; and grant a badge of quality or "EQIPD compliant research" to the user to demonstrate the rigor and integrity of the data collection, analysis, and presentation. Granting such a seal of quality may be beneficial to the research community overall: it would help to improve the openness of science (Kidwell *et al*, 2016; Rowhani-Farid *et al*, 2020) and it may be applied by funding agencies to scrutinize applicants' scientific standards. Of note, this proposal is in line with the *Scoping Report* from the European Commission proposing

that schemes for certification could be one remedy to increase data reproducibility (https://op.europa.eu/en/publication-detail/-/publication/6bc538ad-344f-11eb-b27b-01aa75ed71a1).

> " *. . . the endorsement of a recognition system at different levels of the research ecosystem may provide an effective method to improve research integrity.* "

However, the use of a badge may raise controversial issues. For example, data collected in a CF often represent only a fraction of the results reported in a peer-reviewed paper and a CF badge may mislead readers to believing that best practices were applied to all experiments reported. Nonetheless, the endorsement of a recognition system at different levels of the research ecosystem may provide an effective method to improve research integrity. The EQIPD initiative (http://www.eqipd.online/) is moving in this direction by supplying a Quality System that is shared among the academic and private sectors, as well as funders and publishers. If it were recognized, an EQIPD badge would be valuable to researchers and CFs would play a vital role in supporting high-quality standards in the academic setting.

Data availability

Acknowledgments

Conflict of interest

BG and AB are employees and shareholders at PAASP GmbH.

References

Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533: 452–454

Bespalov A, Bernard R, Gilis A, Gerlach B, Guillén J, Castagné V, Lefevre IA, Ducrey F, Monk L, Bongiovanni S *et al* (2021) Introduction to the EQIPD quality system. *Elife* 10: e63294

Bespalov A, Steckler T (2018) Lacking quality in research: Is behavioral neuroscience affected more than other areas of biomedical science? *J Neurosci Methods* 300: 4–9

Bespalov A, Steckler T, Skolnick P (2019) Be positive about negatives-recommendations for the publication of negative (or null) results. *Eur Neuropsychopharmacol* 29: 1312–1320

Bikovski L, Robinson L, Konradsson-Geuken A, Kullander K, Viereckel T, Winberg S, Roman E, Tsoory M (2020) Lessons, insights and newly developed tools emerging from behavioral phenotyping core facilities. *J Neurosci Methods* 334: 108597

Festing MF (2013) We are not born knowing how to design and analyse scientific experiments. *Altern Lab Anim* 41: P19–P21

Gould J (2015) Core facilities: shared support. *Nature* 519: 495–496

Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG (2016) Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol* 31: 337–350

Heddleston JM, Aaron JS, Khuon S, Chew TL (2021) A guide to accurate reporting in digital image acquisition – can anyone replicate your microscopy data? *J Cell Sci* 134: jcs254144

Kidwell MC, Lazarević LB, Baranski E, Hardwicke TE, Piechowski S, Falkenberg L-S, Kennett C, Slowik A, Sonnleitner C, Hess-Holden C *et al* (2016) Badges to acknowledge open practices: a simple, low-cost, effective method for increasing transparency. *PLoS Biol* 14: e1002456

Knudtson KL, Carnahan RH, Hegstad-Davies RL, Fisher NC, Hicks B, Lopez PA, Meyn SM, Mische SM, Weis-Garcia F, White LD *et al* (2019) Survey on scientific shared resource rigor and reproducibility. *J Biomol Tech* 30: 36–44

Kos-Braun IC, Gerlach B, Pitzer C (2020) A survey of research quality in core facilities. *Elife* 9: e62212

Lazic SE (2016) *Experimental design for laboratory biologists: maximising information and improving reproducibility*. Cambridge, UK: Cambridge University Press

Nuzzo R (2014) Scientific method: statistical errors. *Nature* 506: 150–152

Rowhani-Farid A, Aldcroft A, Barnett AG (2020) Did awarding badges increase data sharing in BMJ Open? A randomized controlled trial. *R Soc Open Sci* 7: 191818

Sherman D (2018) Scientific ethics: what are our responsibilities as facility managers? M&M 2001 expert's session on core facility management. *Microscopy Today* 10: 8–17

Turkiewicz A, Luta G, Hughes HV, Ranstam J (2018) Statistical mistakes and how to avoid them – lessons learned from the reproducibility crisis. *Osteoarthritis Cartilage* 26: 1409–1411