

**RSS RESPONSE TO SCIENCE AND TECHNOLOGY SELECT COMMITTEE INQUIRY ON REPRODUCIBILITY
AND RESEARCH INTEGRITY**

30 September 2021

1. Introduction

1.1. There is an important role for statisticians in understanding and addressing the reproducibility crisis. First, there are a number of ways in which statistical evidence can be distorted as it passes through the research process – from originators, through funders and publishers and eventually to the public. Second, statisticians can play a key role in improving the quality of data, statistical methodology and experimental design throughout this process and improving the ability of people at each stage to assess the quality of methodology and robustness of evidence collected.

1.2. As such, our submission focuses on some of the possible causes of the claimed crisis and looks at what can be done – especially by UK Research and Innovation (UKRI) – to address these problems. We make the following recommendations:

- Recommendation 1* UKRI should consider the feasibility and cost of a UKRI-wide data-driven research methods service and its likely benefits, drawing on evidence of the efficacy of existing similar initiatives.
- Recommendation 2* All projects should specify which member(s) of the research team is/are responsible for both methodology and analysis and the skillsets that enable this – with stronger tests for the presence of adequate skills within research teams.
- Recommendation 3* All UKRI entities should have common standards for the specification of methods, sample sizes and data analysis provided as part of cases for funding. While there may be some discipline-specific requirements these should be consistent with and built around a common core.
- Recommendation 4* When making funding decisions as part of the normal grant-making process, a panel should include at least one member who is expected to judge whether the reviews for a particular proposal have provided a sufficiently robust assessment of methods in it.
- Recommendation 5* When assessing data-driven research proposals in response to a specific funding call, panels should require the presence of a member with expertise in data collection and analysis
- Recommendation 6* UKRI should maintain its commitment to fund new cohorts of doctoral students in advanced data analysis, and the momentum it has built up.
- Recommendation 7* UKRI should evaluate the extent to which its state expectations around doctoral training are met in practice.
- Recommendation 8* UKRI should strengthen the framework for proposal assessment, to ensure that only those able to demonstrate the presence of the necessary skills are able to access funding.
- Recommendation 9* Any data and design associated with UKRI support should be gathered, managed, described and shared – subject to the need to protect personal data – in a way that ensures that any interested and competent party can readily reproduce the results.
- Recommendation 10* UKRI should specify and delineate the role of ethics committees in its decision-making processes, to clarify that the technical review of methods, design, data collection and analysis (while having ethical aspects) is primarily a matter for UKRI.
- Recommendation 11* UKRI should provide additional support for statistical research that is directed towards issues around reproducibility. Statistics as a discipline is well-placed to shed light on the nature of the challenge and to identify tools and techniques to improve the situation.

2. Issues in academia that have led to the reproducibility crisis

- 2.1. Though the idea of a reproducibility crisis is now quite widespread, it is worth noting that the extent of the “crisis” is contested.¹ As the call for evidence notes, Ioannidis’s 2005 paper [Why Most Published Research Findings Are False](#) is perhaps the clearest articulation of the claim that there is a reproducibility crisis in data-driven research.² This paper was based on modelling, rather than empirical evidence – his argument was that reasonable assumptions about the design of studies, biases in conduct, selection in reporting and the small proportion of truly alternative hypotheses investigated meant a high rate of ‘false discoveries’. In a later paper³ he puts the figure at around 50%. By contrast, Jager and Leek⁴ have argued that the false discovery rate is more likely to be around 14%.
- 2.2. Some of the disagreement will relate to the different fields of study. Ioannidis’s work focuses on psychology and neuroscience journals, while Jager and Leek examined top medical journals. We would expect the work in these journals – with randomised control trials and meta-analyses – to be more reliable than first claims of discoveries in psychology and neuroscience. And, given the context, it is perhaps reasonable to think that 14% is quite a high false discovery rate. We do not have a particular view on where the problem is greatest – the point here is just that it is bound vary between different fields.⁵
- 2.3. Alternatively the problem can be tackled empirically – by attempting to replicate the experiments behind past published claims. The Open Science Collaboration [sought to replicate 100 psychology studies](#) finding that whereas 97% of the original studies had statistically significant results, only 36% of the replications did. While this could be viewed as meaning that the majority were false discoveries, the situation is not quite as bad as those figures make it seem. It has also been pointed out that 77% of the new results lay within the 95% predictive interval from the original study. It is perhaps best not to think about a connection between significance and whether something constitutes a “discovery” – but rather to look at the size of the effects. The Open Science Collaboration’s work suggests that replicated experiments on average produced results in the same direction as the original experiments, but were around half their magnitude. So there was a clear bias in the experiments they looked at – but that doesn’t easily lend itself to asserting the percentage of false discoveries.
- 2.4. It is important to diagnose the problem correctly to understand its causes. It seems likely that we are not dealing with deliberate dishonesty – but rather with researchers making particular choices throughout the research process in response to data. These might seem relatively innocuous – eg, selecting which measures to emphasise or how to categorise continuous quantities – and researchers are likely unaware that they constitute questionable research practices. Without really being aware of the consequences of these decisions – and with pressure to get results and publish – these decisions can easily lead to the publication of exaggerated effects
- 2.5. There is good evidence that researchers make these types of decisions. In a [2012 survey of academic psychologists working in the USA](#), a considerable percentage of academics admitted to this type of questionable research practice. The results are shown in Fig.1.

¹ The analysis here summarises §§2-4 of David Spiegelhalter’s 2017 RSS Presidential address, [Trust in numbers](#).

² We use data-driven research to describe activities in which the creation, analysis and representation of data are a central and essential part of a research process, and in which one or more of them forms part of the creative, insightful step that distinguishes research from simple measurement and reporting.

³ Ioannidis (2014), [Discussion: Why “An estimate of the science-wise false discovery rate and application to the top medical literature” is false](#)

⁴ Jager and Leek (2014), [An estimate of the science-wise false discovery rate and application to the top medical literature](#)

⁵ The US National Academies of Science’s report, [Reproducibility and Replicability in Science](#) (2019), is helpful on the difficulties of comparing different fields.

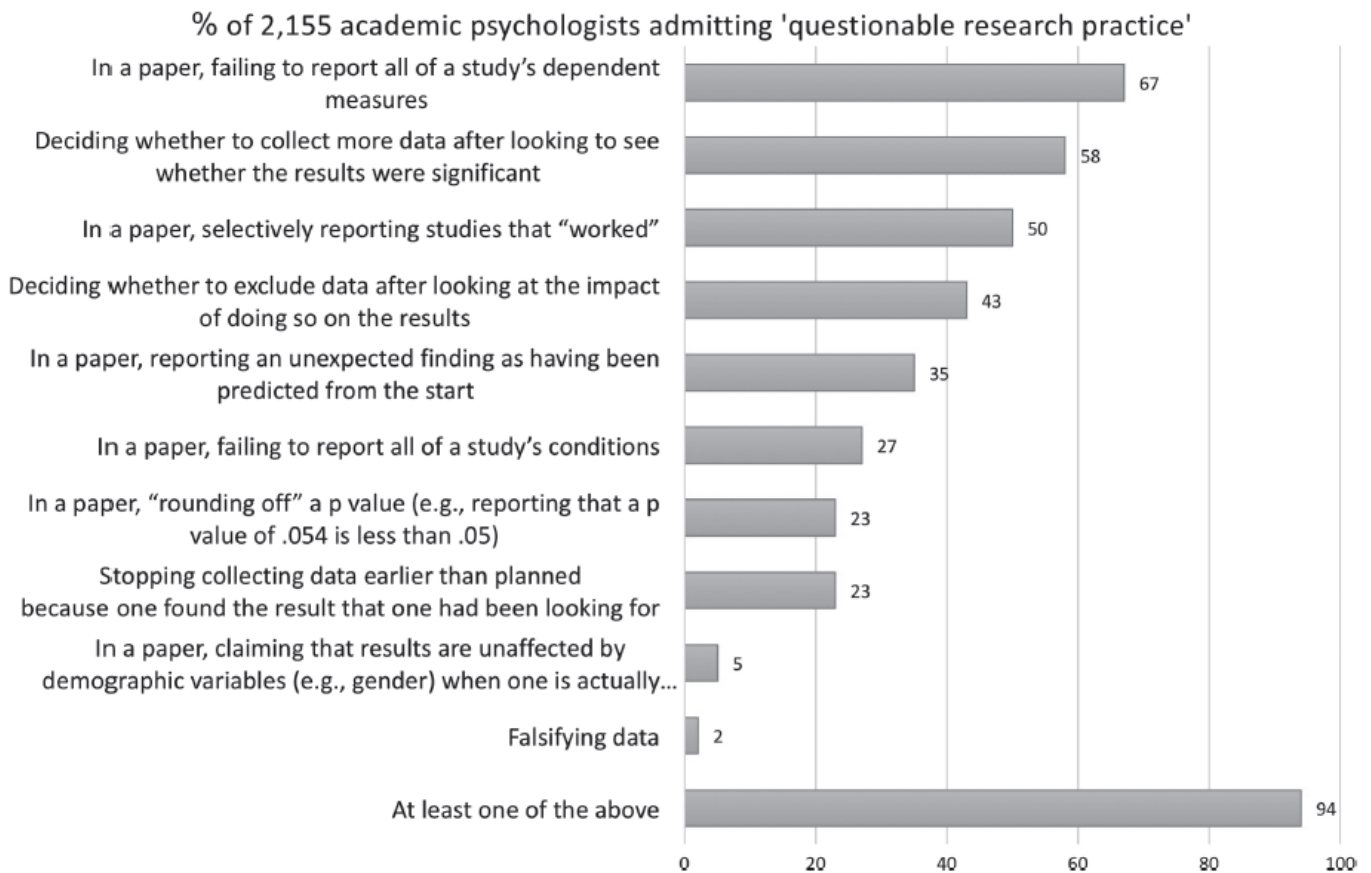


Fig. 1. [Questionable research practices admitted by 2155 US academic psychologists](#), reproduced in [Trust in Numbers](#)

2.6. More concerning than the number of academics who admit to some of these practices is the number that think them to be acceptable. For instance the 50% who admitted selectively reporting studies gave an average score of 1.66 when asked whether this practice was defensible, where 0 = no, 1 = possibly and 2 = yes.

2.7. This survey contained quite a limited set of questionable practices relating to experimentation. If we consider other types of studies – such as surveys – there are many other types of decisions that can be made that might introduce biases: a sample may be chosen because it is the most convenient rather than the most appropriate; questions might be phrased in a leading or misleading manner; or too small a sample might be used. There are plenty of opportunities for people to make poor decisions on the basis that they seem more likely to lead to publishable results.

2.8. It is hard to know what people's motivations are but, if this diagnosis of the problem is right, it suggests that the types of actions that are needed to improve the situation are those that:

- Build awareness in the research community of the impact of methodological choices
- Improve training in statistical methodology
- Ensure that the design of research projects at the initial stage is sound
- Encourage transparency about methodology

In the next section we set out how UKRI, as the main single funder of UK research, can help to improve the situation along these lines.

3. The role of research funders in addressing the reproducibility crisis

3.1. There are a number of stakeholders who have a role in addressing the reproducibility crisis, primarily: funders, publishers and research institutions. Our evidence focuses on what funders – and especially UKRI



– can do to address the problem: this is both because funders of research can be highly influential in this and because – as the Committee notes – UKRI has already been charged with establishing a research integrity committee and is a natural organisation to look at issues around reproducibility. We have identified four specific areas where UKRI can intervene to address issues of reproducibility in data-driven research:

- 3.2. Application stage: Even the most finely-honed and balanced research project will fail if the ideas or assumptions on which it is based, or if the methods and designs it proposes, are wrong. A project can only be as good as the plans and capabilities of the people carrying it out. Deficiencies in approach to data-driven research can appear very early in the process of project formulation and UKRI can take action to help address this.
- 3.3. Appraisal process: UKRI operates a process of peer-review to assess the quality of research proposals. At this stage there is an opportunity to assess the statistical and data-driven aspects of a proposal, however implementing a check here would be complex – it would require a large number of researchers with relevant expertise to take on a substantial additional quantity of work.
- 3.4. The pipeline: As discussed in the previous section, part of the cause of the “crisis” is a deficit in data-driven research skills – both in the UK and globally. UKRI is uniquely placed to bolster these skills among doctoral students.
- 3.5. Research integrity and ethics: The evidence suggests that there are some issues around research integrity that may contribute to issues with reproducibility. UKRI, while the largest single funder of UK research and innovation, cannot hope to change the landscape on its own. While UKRI has a limited number of levers in this area, it can at least seek to ensure that the research it funds abides by the highest standards of research integrity.
- 3.6. We propose a number of concrete steps that UKRI can take under each of these four areas.⁶

a) Application Stage

- 3.7. Some of the issues that have led to the reproducibility crisis can be embedded at the earliest stage of project development. If the methodologies or experimental designs proposed are unsound then a research project is more likely to produce evidence that cannot be reproduced. There is a clear need for research teams who are applying to UKRI to have access to the best-possible advice regarding data-driven research – and there are a number of ways that UKRI could facilitate this.
- 3.8. One potential approach would be for UKRI to provide data-related advice and resources to be used by researchers who do not have the necessary skills themselves. At the most basic level, a repository of methods could be developed and made available to all researchers, but additional and more involved systems of advice and engagement could be conceived. The possibilities are wide-ranging, and the repository solution is tempting. But even a basic repository would be a significant undertaking, both in terms of development and maintenance. **We believe that UKRI should consider the feasibility and cost of a UKRI-wide service to advise on data-driven research methods.** It is not clear to us how successful this would be – not least because it would rely on researchers consulting the service when they either do not know they need to or simply do not want to. However, other organisations – eg, the National Institute for Health Research – run this type of [research design service](#) and as part of this review, UKRI should assess how effective similar services are.
- 3.9. In some areas – notably biomedical sciences – it is routine to identify a statistically competent member of each research team with responsibility for methodology and analysis. This seems a useful way of ensuring that research projects are based upon sound statistical methodology and experimental design. The

⁶ These proposals are based on a joint UKRI-RSS review of statistics and data-driven research across the research councils. The report was produced as an internal UKRI document, authored by: Deborah Ashby (then RSS President); Frances Burstow (then Deputy Director Skills ESRC); Sir Ian Diamond (then UKRI Board); Guy Nason (then Vice President Academic Affairs RSS); Jennifer Rubin (then Executive Chair ESRC); Hetan Shah (then Executive Director, RSS); Alex Hulkes (then Strategic Lead Insights, ESRC).



practice is not, however very widespread. UKRI should require, across all areas of its work, the inclusion in the teams it supports of sufficient data-related expertise both at the inception of projects and during their delivery. **All projects should specify which member(s) of the research team is/are responsible for both methodology and analysis and the skillsets that enable this. This would also require stronger and more specific tests for the presence of adequate skills within research teams during the review process and perhaps a requirement that statistical skills are backed by qualifications or accreditation such as the RSS's Chartered Statistician.**

3.10. When submitting an application to a research council, research teams set out – as best they can – their proposed methodology. This is not always easy: the project's pathway will often be uncertain and it is not always clear how much importance is placed on various requirements such as experimental design, attention to sample sizes and the justification for overall methodology. This can have the unwanted effect of leading applicants to conclude that too much detail could be counterproductive (especially when the guidance is restrictive on the types of methodological approaches that are fundable). We believe that the current flexibility and latitude in approaches to describing data-driven research in UKRI funding proposals is potentially counterproductive. **All UKRI entities should have common standards for the specification of methods, sample sizes and data analysis provided as part of cases for funding. While there may be some discipline-specific requirements these should be consistent with and built around a common core.**

3.11. To achieve this aim UKRI should review the guidance provided in relation to structuring an application and consider separating out issues of experimental design, methodological choice, statistical analysis and other key data-driven research issues into a standalone technical annex which sits outside the normal page limit. This would allow the data aspects of a project to be assessed more directly and it would remove painful trade-off experienced by applicants, albeit at the expense of longer proposals and the problems that might entail.

Recommendation 1: UKRI should consider the feasibility and cost of a UKRI-wide data-driven research methods service and its likely benefits, drawing on evidence of the efficacy of existing similar initiatives.

Recommendation 2: All projects should specify which member(s) of the research team is/are responsible for both methodology and analysis and the skillsets that enable this – with stronger tests for the presence of adequate skills within research teams.

Recommendation 3: All UKRI entities should have common standards for the specification of methods, sample sizes and data analysis provided as part of cases for funding. While there may be some discipline-specific requirements these should be consistent with and built around a common core.

b) The appraisal process

3.12. After applications have been submitted to UKRI, they are assessed in a range of ways falling under the umbrella term “peer review”. There is considerable variability in the description and implementation of appraisal processes across UKRI, including the treatment of data-driven research. Currently none of the nine research councils of UKRI systematically uses reviewers with statistical or data expertise to assess the statistical or data-driven aspects of a research proposal. In some areas – eg, High Performance Computing, which requires a technical assessment before a proposal is submitted – this type of process is in place and it is tempting to think that there should be a specific stage in the review process to assess statistical methodology and experimental design.

3.13. There would be some clear advantages to this approach. Most notably it would create a clear and verifiable check in the system and it would also signal the importance of these issues to applicants. However, there is a limited pool of researchers who would be able to conduct this work and it would place a very heavy workload on those individuals. Furthermore it assumes that there is a single correct approach and sets up a single reviewer as the arbiter thereof, which risks stymying developments in emerging areas where the best approach is not at all clear cut. So we do not think that this approach is an effective way of tackling the problem.

- 3.14. An alternative is to look at the process by which research councils select reviewers and evaluators. Not much is known about how reviewers and evaluators are assigned to review proposals – in general it seems as though a composite approach to the task of securing sufficient comment is used, one that seeks to cover all aspects of a proposal across a number of individual reviewers. But these aspects can be intersecting and multidimensional, and they all need to be met within the expertise of (on average) three or four reviews for a proposal. It is difficult for us as an external organisation to make specific proposals for this type of complex process, but UKRI should assess whether there are ways to bring data-informed expertise into the process.
- 3.15. After the initial review phase, research councils typically have a panel phase – at which a panel of experts make a funding recommendation. There are a number of process improvements that UKRI could introduce with a view to improving the quality of statistical and data-driven research methodology in funded projects.
- 3.16. Moderating panels are defined by the requirement that additional technical information and opinion is not introduced into the discussion – they are intended to use only the assessments from the first stage of reviews to avoid the risk that applicants are affected by comments without having a chance to respond to them. This does not preclude the possibility that **a moderating panel should include at least one member who is expected to judge whether the reviews for a particular proposal have provided a sufficiently robust assessment of methods in it.** It is established practice that inadequately reviewed proposals are deferred to later panels pending additional reviews, and such a change would merely formalise and specify this requirement and practice in relation to data-driven research. In cases where a decision is delayed, it is important that a robust assessment of the methods is sought as quickly as possible so that researchers are not left waiting on a decision for too long because of a failure of the review process. Research councils should publish details of the percentage of decisions that are delayed for this reason.
- 3.17. Panels convened for specific calls for proposals tend not to have the option of deferral for further reviews, potentially creating a problem: if a decision must be made at a specific meeting, the panel can only use the reviews already available. But these panels also tend to have elements of the assessing panel process, smoothing the transition from one mode of working to the other. **Assessing panels should require the presence of a member with expertise in data collection and analysis.** The exact details of the role would necessarily vary from panel to panel, as currently do the roles of all other panel members.

Recommendation 4: When making funding decisions as part of the normal grant-making process, a panel should include at least one member who is expected to judge whether the reviews for a particular proposal have provided a sufficiently robust assessment of methods in it.

Recommendation 5: When assessing data-driven research proposals in response to a specific funding call, panels should require the presence of a member with expertise in data collection and analysis.

c) The pipeline

- 3.18. It is widely accepted that there is a deficit in data-driven research skills in the UK and globally.⁷ The deficit is not uniform across all UKRI research, because the requirement for skills is not uniform. Every population of researchers will have its own distribution of skills, with some disciplines being in general more proficient than others either of necessity or by chance. -
- 3.19. UKRI develops the next generation of new researchers and **UKRI should maintain its commitment to fund new cohorts of doctoral students in advanced data analysis, and the momentum it has built up.** These students could be linked to projects or based in major data analysis centres. It is important to

⁷ The Royal Society's Dynamics of Data Science Skills report sets this out clearly. For an example of how this plays out in the academic community see [this survey of misconceptions relating to the basic and well-established technique of linear regression in published literature.](#)



stress that statistical and data analytical skills benefit researchers in a wide range of disciplines and UKRI should ensure that this training is widespread.⁸

3.20. Providing the funding for this training is an important step, but UKRI must also ensure that the training is adequate – that all students receiving the training can make decisions around methodology and design, understand how to manipulate data and report data analyses. **To support this aim UKRI should evaluate the extent to which its stated expectations around doctoral training are met in practice. A properly designed survey of sufficient size, using appropriate methods, should be carried out to ascertain both the prevalence of issues and their nature across the research councils.** Such a survey would also help to establish just how many data-skilled students UKRI supports and what the prevalence of data-related skills is across the student landscape.

3.21. UKRI could also usefully encourage existing researchers to develop their analytical skills. **The most useful intervention UKRI might make would be to strengthen the framework for proposal assessment, to ensure that only those able to demonstrate the presence of the necessary skills are able to access UKRI funding.** UKRI could also encourage existing researchers to develop their analytic skills through funding streams that promote the secondary analysis of existing data sets accompanied by training to upskill the applicants.

Recommendation 6: UKRI should maintain its commitment to fund new cohorts of doctoral students in advanced data analysis, and the momentum it has built up.

Recommendation 7: UKRI should evaluate the extent to which its stated expectations around doctoral training are met in practice.

Recommendation 8: UKRI should strengthen the framework for proposal assessment, to ensure that only those able to demonstrate the presence of the necessary skills are able to access funding.

d) Research integrity and ethics

3.22. Two-thirds of the UK's gross expenditure on R&D (GERD) occurs in the business sector; by contrast, UKRI is probably responsible for allocating about one sixth of GERD. While UKRI is the largest single source of funding, its levers are limited in the context of the wider funding ecosystem. However, UKRI still has some levers and it should use these as much as possible while also collaborating with other major stakeholders.

3.23. First, **UKRI should provide additional support for statistical research that is directed towards issues around reproducibility. Statistics as a discipline is well-placed to shed light on the nature of the challenge and to identify tools and techniques to improve the situation.**

3.24. UKRI can set expectations for standards of research integrity for the projects that it funds: **its processes should operate on the expectation that data and design associated with UKRI support should be gathered, managed, described and shared – subject to the need to protect personal data – in a way that ensures that any interested and competent party can readily reproduce the results.** This has implications for the accessibility of data, metadata and code needed to implement published analysis. While there may be some discipline-specific differences in implementation such a function should be built around a common core of aims and characteristics that are consistent across research areas. There is some confusion over data access in the context of GDPR:⁹ UKRI's guidance is not currently strong enough on this and there should be an expectation that UKRI actively enables greater data sharing between researchers.

⁸ For more detail see our recommendations for the Economic and Social Research Council's doctoral training programmes.

⁹ For a discussion of this see Carrigan's (2019) [Saving lives, or costing them? The unintended consequences of data protection](#)



3.25. If our suggestion in paragraph 3.11. to include a technical annex that details issues around methodological choices is taken up, it may also be beneficial to publish these documents for all funded research – in much the same way that UKRI already shares broader summaries of project objectives and plans for impact. This would then form part of a trail that would encourage and demonstrate reproducibility and good conduct in research.

3.26. The fact that poor research and innovation practice is undoubtedly an ethical issue as well as a practical one creates a situation in which it is possible to argue for a key part of the UKRI decision-making process to be outsourced to another body. If questions of method are viewed solely through the prism of research ethics, it is entirely possible that the case for a project’s methodology and analytical approach will be made to an external ethical review body. We believe that this is something that should be resisted – especially if it would mean that applicants could secure funding with an assertion of approval from an external body in relation to a project’s experimental design or methods. Were that to happen, UKRI would lose control over the criteria and rigour of the process. This is not to say that there shouldn’t be an independent ethical process as well – to prevent unethical research projects from being funded – but it is important that projects are not approved because they are deemed to be ethical rather than because they meet high technical standards.

3.27. All considerations of methodological acceptability should instead be made transparently, with external ethical assurances being sought before funding is approved. **We recommend that UKRI should specify and delineate the role of ethics committees in its decision-making processes, to clarify that the technical review of methods, design, data collection and analysis (while having ethical aspects) is primarily a matter for UKRI.**

Recommendation 9: Any data and design associated with UKRI support should be gathered, managed, described and shared – subject to the need to protect personal data – in a way that ensures that any interested and competent party can readily reproduce the results.

Recommendation 10: UKRI should specify and delineate the role of ethics committees in its decision-making processes, to clarify that the technical review of methods, design, data collection and analysis (while having ethical aspects) is primarily a matter for UKRI.

Recommendation 11: UKRI should provide additional support for statistical research that is directed towards issues around reproducibility. Statistics as a discipline is well-placed to shed light on the nature of the challenge and to identify tools and techniques to improve the situation.

4. Closing remark: a role for a national committee on research integrity

4.1. We have sought to set out the types of actions that are needed to address challenges around reproducibility and have made a number of recommendations setting out how UKRI can: build awareness in the research community around the importance of methodological choices; improve training in statistical methodology; ensure the soundness of the design of research projects; and encourage transparency around methodology.

4.2. The proposed introduction of a national committee on research integrity is welcome. Such a body could usefully serve to provide oversight of the implementation of any actions that it is agreed that UKRI should take in order to improve reproducibility of data-driven research. There would also potentially be a key role for this type of organisation in the context of our final recommendation – as the national committee could be a useful way to feed in ethical considerations into decision-making while still ensuring the role of UKRI as final decision-maker on the technical merits of proposals.

