

## Improving research practices in energy: practical guidance for greater transparency, reproducibility, and quality

Gesche M. Huebner<sup>1,2\*</sup>, Michael J. Fell<sup>1</sup>, Nicole Watson<sup>1</sup>

<sup>1</sup> UCL Energy Institute, University College London, Gower Street, London, UK.

<sup>2</sup> UCL Institute for Environmental Design and Engineering, University College London, Gower Street, London, UK.

\* corresponding author: g.huebner@ucl.ac.uk

**PLEASE NOTE: This manuscript has not been peer reviewed.**

**Date of upload 20 July 2020**

### Abstract

Energy use is of crucial importance for the global challenge of climate change but also an essential part of daily life. Hence, research on energy needs to be robust and valid. Other scientific disciplines have experienced a reproducibility crisis, that is, existing findings could not be reproduced in new studies, and energy research might be impacted as well. In this paper, we suggest the 'TReQ' approach to improve the research practices in the energy field and arrive at greater *Transparency*, *Reproducibility*, and *Quality*. We acknowledge the specific challenges of energy research and suggest a highly adaptable suite of tools that can be applied to research approaches across this multi-disciplinary and fast-changing field. In particular, we introduce preregistration of studies, making data and code publicly available, using preprints, and employing reporting guidelines to heighten the standard of research practices within the energy field. We argue that through wider adoption of these tools, we will be able to have greater trust in the findings of research used to inform evidence-based policy and practice in the energy field.

**Practice relevance:**

The paper gives concrete suggestions on how and when to use preregistration, open data and code, preprints, and reporting guidelines, offering practical guidance for energy researchers for improving the **T**ransparency, **R**eproducibility, and **Q**uality of their research. We show how employing TReQ tools at appropriate stages of the research process can assure end-users of the research that best practices were followed. This will not only increase trust in research findings, but can also deliver other co-benefits for researchers, such as more efficient processes and a more collaborative and open research culture. We demonstrate how employing the TReQ tools can help remove barriers to accessing research both within and outside of academia, improving the visibility and impact of research findings. Finally, we present a checklist that can be added to publications to show how the tools were used.

**Keywords:**

Energy, transparency, quality, reproducibility, open science, best practice

## 1. Introduction

Energy use is key to global challenges such as climate change, while also playing an important role in our daily lives. But how sure can we be that research in this area is providing reliable findings? In this paper, we argue that limited employment of principles and tools to ensure transparency and quality could be making this hard to assess.

A degree of openness is widely viewed as an essential component of good research. Unless sufficient details of studies are shared, it can be hard for other researchers to tell if conclusions are justified, check the reproducibility of findings, and undertake effective synthesis of evidence. However, it is not self-evident which aspects of studies need to be shared. A range of tools and practices have been developed to guide researchers on what to share, and when. These include guidelines on which details of studies to report, pre-registration of theory-testing work, and the sharing of data and code. While such approaches are now widely employed in some disciplines – such as medicine and psychology – in energy research they remain largely a niche concern (Huebner et al. 2017). Consequently, evidence-based policy and practice may be built on shaky foundations.

In this paper, we provide a practical guide to the tools and principles we believe energy researchers should now be bringing to their work. Specifically, we cover preregistration of studies/analysis, use of reporting guidelines, sharing of data and analysis code, and publishing preprints. We recognise that real-world research is often messy and unpredictable and suggest why and how researchers should use the tools despite (or even because of) this. We suggest that the use of these tools can improve the transparency, reproducibility (where appropriate), and quality (TReQ for short) across the field.

### 1.1. General benefits of TReQ

Transparency has a range of meanings (Ball 2009); we take it to mean the accessible sharing of research processes, findings, and data. Transparency is at the core of the

approach because it enables informed interpretation and synthesis of findings, as well as supporting reproducibility and general research quality (Miguel et al. 2014).

Reproducibility means that independent studies testing the same thing should obtain broadly the same results (Munafò et al. 2017), for studies where there is the assumption that the findings hold beyond the original sample tested<sup>1</sup>. When best research practices are followed, if the study is reproduced, the same conclusions should be reached.

Some disciplines are finding themselves in a reproducibility crisis, an ongoing scientific crisis that indicates that the results of many scientific studies are difficult or impossible to reproduce. When 100 studies from three high-ranking psychology journals were repeated, only 36% of the repeats had significant findings, compared to 97% of the original studies, with a mean effect size of about half that of the original studies (Open Science Collaboration 2015). Preclinical research also showed a spectacularly low rate of successful reproduction of earlier studies (Prinz, Schlange and Asadullah 2011; Begley and Ellis, 2012). In a survey, 70% of approximately 1500 scientists indicated that they have tried and failed to reproduce another scientist's experiments, and more than half had failed to reproduce their own experiments (Baker and Penny 2016). This is a problem if policy and research are being built on effects which are either not real, or do not apply in the context in which they are being deployed.

In rare cases, outright data fraud happens (Stroebe, Postmes and Spears, 2012), which the TReQ approach discourages or makes easier to detect. More importantly, it can deter so-called questionable research practices (QRP), or “design, analytic, or reporting practices that have been questioned because of the potential for the practice to be employed with the purpose of presenting biased evidence in favour of an assertion” (Banks, O’Boyle, et al. 2016, p. 3).

---

<sup>1</sup> We focus on reproducibility in the sense outlined above; different to replicability which we, in line with Munafò et al. (2017), mean to rerun the same analysis on the same data.

## Improving research practices in energy (preprint 20 July 2020)

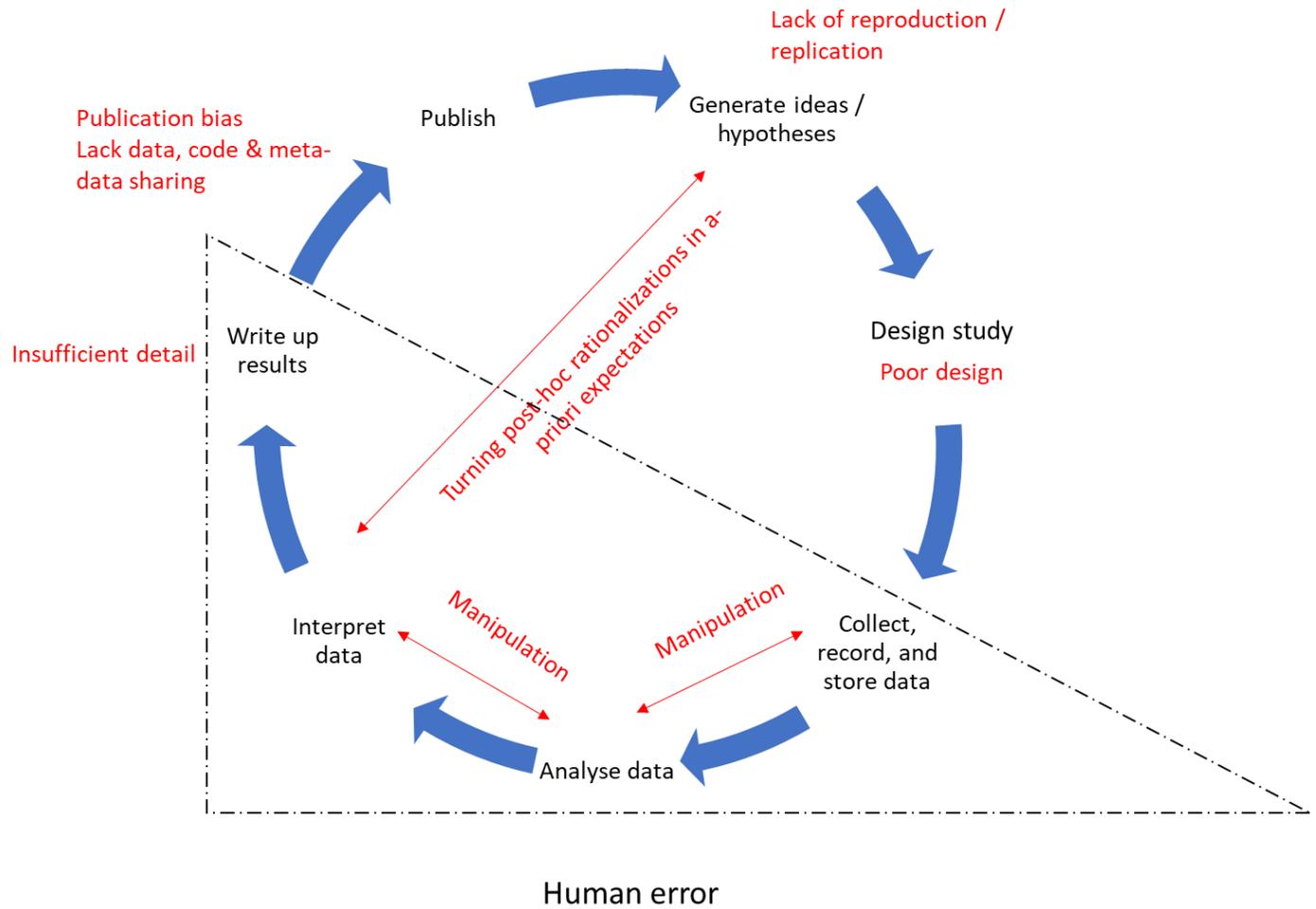


Figure 1 (adapted from (Chambers et al. 2014)) shows a circle map of scientific process and the points where issues that threaten transparency, reproducibility and quality of research can come into play.

Moving clockwise from the top, we highlight that pure reproducibility studies are rare as they are often viewed as having a lower status than novel discoveries (see Park 2004). Poor study design can encompass a wide range of issues, including missing important covariates or potentially biasing factors, or an increased likelihood of false positives or false negatives (Button et al. 2013).

Manipulation as a generic term stands for changing, analysing, and reporting data in ways that support outcomes deemed favourable by the researcher. It is important to stress that QRPs can be employed quite innocently without any intent to mislead, for

instance, through human error (see dashed triangle, Figure 1). Regardless of intent, the impact is to mislead nonetheless. Both qualitative and quantitative work are open to manipulation of data collection or analysis to produce interesting or 'significant' findings. Examples of the latter include only reporting part of the results, rounding down p-values, and p-hacking, or taking arbitrary decisions with the data in order to achieve results below the typical significance threshold of a p-value  $< 0.05$  (Banks, Rogelberg, et al. 2016; Simmons et al. 2011). The impact of p-hacking is to overstate the strength or even existence of associations or effects.

One study shows that more than 70% of respondents indicated they had published work not reporting all dependent measures (John, Loewenstein and Prelec 2012). Taking liberal decisions around p values was reported by 11% of researchers and turning post-hoc explanations into a-priori expectations by about 50% (Banks, O'Boyle, et al. 2016).

Another form of manipulation applicable primarily in the domain of quantitative research is HARKing, or 'hypothesizing after results are known'. For example, a researcher may find a significant association between two variables and then write up the study as if they anticipated the association. This portrays an exploratory study as a confirmatory one (Kerr 1998). Again, this is often prompted by a perception, that journals favour studies with significant effects to those that report null results (Dickersin et al. 1987). This 'publication bias' creates a skewed representation of true findings in the published literature. Withholding data, code, meta-data, or detail on important characteristics of studies makes it harder to spot such practices, and less likely that high quality replication studies can be carried out.

Using tools to increase transparency, reproducibility, and quality also have wider benefits around the accessibility of research. A substantial amount of funded research is paid for by taxpayers through government funding (ONS 2020), yet many research papers are not publicly available, although the proportion is decreasing (Piwowar et al. 2018). The inaccessibility of research reports has a general slowing and discriminatory effect on the progress of research (Suber 2013; Eysenbach 2006). Open access, on the other hand, can support a range of social and economic benefits (Tennant et al. 2016; Fell 2019). As such, funders are increasingly mandating open access publishing.

However, routes for open access include depositing the manuscript in a repository where it is embargoed for a certain period; hence only accessible after a delay (e.g. UKRI n.d). TReQ tools contribute to making science openly available.

## 1.2. The importance of TReQ in energy research

Whilst more transparent, reproducible and high-quality research is important in any area, it is perhaps especially pressing in the applied area of energy research. Climate change is indubitably the biggest challenge society is facing. Much research in the energy field is directly relevant to climate change mitigation or adaptation. In order to inform policy, evidence needs to be high quality and reliable. Yet, despite this crucial importance, energy research remains behind other disciplines when it comes to best research practices (Huebner et al. 2017; Pfenninger et al. 2018).

This is likely the case for a number of reasons. Energy research is highly multidisciplinary and uses a multitude of methods, such as interviews, focus groups, surveys, field and lab experiments, case studies, monitoring and modelling. This diversity of academic background and methods used makes it harder to design gold standards applicable to the majority of researchers, and for researchers to judge the quality of work outside their discipline (Eisner 2018; Schloss 2018). In a more homogenous field, agreement on best practices will be easier to find.

In more 'pure' disciplinary areas such as psychology or physics, a primary research interest is often in establishing general and enduring principles. In energy, however, research is often much more focused on the current situation with a strong emphasis on contextual factors, fully aware that ten years down the line, things will have changed.

Related to this, there are a number of research areas within energy where reproducibility, in a strict sense, might not even be an appropriate term to discuss. Many qualitative and participatory research projects focus on specific case studies where with different participants and results are expected to differ. Nevertheless, a degree of generalizability is always important if research evidence is to be of use outside of the direct context of a research study.

A final practical factor that makes it harder to reproduce previous research is that (especially) field trials are extremely time- and money-consuming, more than in

many other disciplines. Similarly, to build an energy system model of a country easily takes several years. These are often one-shot undertakings, with no reasonable expectation that they could later be recreated to check the results.

Despite all these challenges, we argue that energy researchers can -- and should -- be doing much more to integrate TReQ research best practice into their work. We suggest that far from presenting an additional burden of work, it can bring co-benefits. In the remainder of this paper we present a set of simple tools which almost all energy researchers should now consider employing.

## 2. The TReQ approach

From the suite of approaches that support TReQ research, we have chosen to focus on four: study preregistration; reporting guidelines; preprints; and code/data sharing. Our criteria for deciding which tools to focus on was that they should all be:

- Applicable to the wide multidisciplinary variety of research approaches employed in energy.
- Flexible in terms of how they can be employed, so that researchers can use them in ways that they find most useful rather than feeling constrained by them.
- Low barrier to entry – they are easy to pick up and require little specialist knowledge (at least in basic applications).

### 2.1. Preregistration of analysis plans

Preregistration involves setting out how researchers plans to carry out and analyse research before undertaking the work, and (when applicable) what you expect to find (Nosek et al. 2018). Preregistration, whilst more common in deductive (theory-testing) quantitative research, can be applied to any type of research including qualitative research (Haven and Van Grootel 2019) and modelling approaches (Crüwell and Evans 2019). In its basic form, for all types of research, preregistration should specify the study aims, the type of data collection, the tools used in the study, and the data analysis approach. For a more quantitative approach, a preregistration usually includes details on the key outcome measures and details on statistical analysis, including how missing

data and outliers will be handled. Where applicable, concrete hypotheses are also listed. Such a preregistration, often called pre-analysis plan (PAP), is put online, with a certified time-stamp and registration number. It can either be immediately shared publicly, or kept private until a later date (such as publication of a paper).

Preregistration has three main benefits. Firstly, it adds credibility to results because a researcher cannot be accused of QRPs, such as changing their analyses and expectations afterwards to fit the data. As discussed, these practices give an impression of greater confidence in results than is warranted, and allow presenting an exploratory finding as a confirmatory one (Wagenmakers and Dutilh 2016; Simmons et al. 2011).

Secondly, preregistration contributes to mitigating publication bias and file drawer problems. Academic publishing is biased towards publishing novel and statistically significant findings to a much greater extent than non-significant effects (Rosenthal 1979; Ferguson and Brannick 2012). If non-significant findings are not published, this reduces scientific efficiency since other researchers might repeat approaches that have been previously tried and failed. Whilst it is unrealistic to expect researchers to evaluate numerous study preregistrations and follow up on unpublished results, a systematic review might do so. Furthermore, a researcher finding a null effect after having written and followed a pre-analysis plan might be more likely to attempt publishing these findings. Registered reports, where a journal reviews the equivalent of a detailed pre-analysis plan and provisionally commits to publishing a paper regardless of the results, show a greater frequency of null-results than conventionally published papers (Warren 2018).

Thirdly, preregistration brings direct benefits to the researcher (Wagenmakers and Dutilh 2016). It helps with planning study details and getting early input into research design and analysis. This will likely lead to better conducted studies and faster analysis after data collection. It also allows putting an early stake to the area you are working in and sends the message that the researcher is committed to research transparency. Whilst preregistration might seem time-consuming, this time is actually regained at the

data analysis stage. Preregistration does not rule out exploratory analysis on the data, as long as any deviations from the pre-analysis plan are noted in publications.

In some disciplines e.g. economics, planned analyses are very complex and preregistrations documents might become unwieldy long if considering all possible options for nested hypotheses and analyses. In such cases, a simplified preregistration laying out key aspects may be deemed sufficient (Olken 2015).

There are multiple platforms for publishing pre-analysis plans. [AsPredicted.org](https://AsPredicted.org) has a simple form that can be applied to a wide range of studies. Once completed, the document is saved in a time-stamped version that can be made publicly available immediately or a later stage. The [Open Science Framework](https://OpenScienceFramework.org) website maintains a useful list of [templates](#), including those specific to qualitative research and secondary data analysis.

At latest, pre-registration should be done before analysing data. However, best practice is to work on the preregistration whilst designing the study to reap the benefits from it. When writing up the research, details of the PAP (e.g. hyperlink and registration number) are included in any output. Deviations from the pre-specification should be noted and justified (see e.g. Nicolson et al. 2017 p.87).

## 2.2. Reporting guidelines

When reporting research, it can be difficult to decide which details are important to include, and which can be left out. Missing out important details of the work creates several issues. It can make it hard for readers to judge how well-founded or generalisable the conclusions are and may make the work difficult or impossible for others to reproduce. Future evidence reviewers may have difficulty integrating the method and findings, limiting the extent to which work can contribute to evidence assessments.

To help address these problems, sets of reporting guidelines have been developed spelling out precisely which details need to be included for different kinds of study. This include questions such as sampling methods and how recruitment was undertaken,

although these naturally vary between study types. Details of some prominent guidelines are provided in Table 1. Further options and a guide to which checklist to use are available at [Equator Network](#) (Equator Network, 2016).

Table 1: Some of the main reporting guidelines likely to be applicable to energy research.

<b>Study type</b>	<b>Reporting guidelines</b>	<b>Notes</b>	<b>Reference</b>
Randomized trial	<u>CONSORT</u>	Requires flowchart of phases of trial and includes 25-item checklist.	Schulz, Altman and Moher 2010
Systematic review	<u>PRISMA</u>	Flowchart and 27-item checklist.	Moher et al. 2009
Predictive model	<u>TRIPOD</u>	22-item checklist.	Collins et al. 2015
Qualitative study (interviews and focus groups)	<u>COREQ</u>	32-item checklist.	Tong, Sainsbury and Craig 2007)

Beyond making research more interpretable and useful to research users, using reporting guidelines can have a range of benefits for researchers themselves. They give added confidence that important details are being reported, and a way of justifying such choices (for example, in response to peer review comments). They can also make it quicker and easier to write up reports, drawing on the guidelines to help structure them. Guidelines are not only useful at the reporting stage. By becoming familiar early on with standard reporting requirements, researchers can ensure they are considering all the details they will need to report and make note of important steps during research design and collection. Checklists can be used to explicitly structure reports or, at the very least, as a check to ensure that all relevant details are included somewhere. It is usual to cite the guidelines that are being followed.

Following reporting guidelines is, however, not always straightforward. They are often developed for a very specific purpose, such as medical randomised control trials. Even if researchers identify the most suitable type of guideline for their own work, the guidelines may still call for reporting of irrelevant details. In such cases we recommend that they are followed at researchers' discretion – but it is better to consider and exclude points than risk missing out reporting on important details.

### 2.3. Preprints of papers

Scholarly communication largely revolves around the publication of articles in peer reviewed journals. The peer review process, while imperfect, fulfils the important role of performing an independent check on the rigour with which the work was conducted and the justifiability of the conclusions that are drawn. However, peer review and other stages of the academic publication process means that substantial delays can be introduced between results being prepared for reporting, and publication. According to [SciRev](#), the average time taken for the review process alone is 17 weeks, and may be substantially longer. Whilst many journals now allow researchers to make available unformatted versions of accepted papers for free (on a 'green' open access basis), this is usually only after an embargo period.

This results in an extended period before potentially useful findings can be acted upon by both other researchers and those outside academia, especially where researchers are unable to afford to pay for open access or journal subscriptions, including universities in developing countries (Suber 2013, p.30). This is especially problematic in areas such as tackling the climate crisis, when rapid action based on the best available evidence is essential.

The response of the academic community to this has been the institution of preprints, or pre-peer review versions of manuscripts that are made freely accessible. While producing preprints has been standard practice in disciplines such as physics for decades, they are still comparatively rare in many areas of energy research. Nonetheless, preprints can bring substantial benefits to both researchers and research users. They allow early access to, and scrutiny of, research findings, with no affordability constraints. They also allow authors to collect input from peers prior to (or in

## Improving research practices in energy (preprint 20 July 2020)

addition to) the peer review process, with the potential to improve the quality of the reporting or interpretation. They also allow researchers to establish precedence of work or findings, get cited earlier, and avoid unnecessary duplication of work across groups. Some have reservations about preprints. Because they are not peer reviewed, there are legitimate concerns that if preprints present findings based on erroneous methods, this could be dangerously misleading for users. Authors may similarly be worried about putting their work out for wider scrutiny without the independent check that peer review provides. While these are valid concerns, part of the rationale for preprints is to serve a similar function to peer review and pick up problems so they can be corrected in later revisions, which then become the (published) version of record. Preprints are (or at least should be) clearly labelled as such, and a 'buyer beware' approach taken on behalf of readers. Indeed, peer review does not obviate the need for critical use. Some are concerned that preprints may be considered by a journal as a 'prior publication', making it harder to publish the work. While it is always important to check, almost all quality journals now explicitly permit preprints. Finally, it is important to set any concerns against the benefits brought by speedier publication described above, especially to for those who cannot pay to publish or access research outputs.

Publishing a preprint is as easy as selecting a preprint server and uploading a manuscript at any point prior to 'official' publication (most commonly around the point of submission to a journal). A range of such servers exist, probably the most well-known of which is arXiv.org, a preprint server for the physical sciences. The lack of a clear preprint server of choice for a multidisciplinary field such as energy is likely to be a contributing factor to their lack of uptake. Relevant options to consider include PeerJ, preprints.org, SocArXiv, and PsyArxiv. It is also possible to share preprints through institutional repositories and more general scientific repositories such as Figshare. Preprint servers will usually assign a digital object identifier (DOI), and preprints will therefore show up in (for example) Google Scholar searches. Because preprints are assigned a DOI, it is not possible to delete them once they have been uploaded. However, revised versions can be uploaded, and once the paper is published in a peer reviewed version, this version can be linked to the preprint.

## 2.4. Open data and open code

Open data means to make data collected in a research study available to others, usually in an online repository. Open Data must be freely available online and in a format that allows other researchers to re-use the data (Huston, Edge and Bernier 2019). Open code is equivalently making the computer (or qualitative analysis) code that was used to analyse certain data publicly available. Both fall under the umbrella term of 'open science'. Open data practices are not as widespread in energy compared to other research areas (Pfenninger et al. 2017; Kjærgaard et al. 2020).

Both open data and open code can allow the identification of errors, both intentional and unintentional. Open data and open code allow exact replication, i.e. re-running of analyses. The real strength of publication of datasets and computer code is in combination with PAPs: without pre-specification of the data collection, cleaning, and analysis processes, uploaded data could still be manipulated, by for example omitting certain variables or deleting outliers. Together with PAPs, cherry-picking of results and other forms of p-hacking could be uncovered.

Open data and open code can increase productivity and collaborations (Pfenninger et al. 2017), as well as the visibility and potential impact of research. An analysis of more than 10,000 studies in the life sciences found that studies with the underlying data available in public repositories received 9% more citations than similar studies for which data were not available (Piwowar and Vision 2013). Research has also indicated that sharing data leads to more research outputs using the data set than when the data set is not shared (Pienta et al. 2010). Open code is also associated with studies being more likely to be cited than those that do not share their code (Vandewalle 2012). Open data can also be used for other purposes, such as education and training.

This increased usage and visibility through open code and open data can advance science further and more quickly, and also brings personal benefits to a researcher, given the importance of scientific citations and impact measures. Sharing data and code is most common in quantitative research, but is possible and encouraged in qualitative research (McGrath and Nilsson 2018).

Conversely, the drive to publish in competitive academic environments may play a role in discouraging open data practices: data collection can be a long and expensive process and researchers might fear premature data sharing may deprive them from the rewards of their effort, including scientific prestige and publication opportunities. However, the benefits on increased visibility, credit and citations should outweigh this perceived concern.

Making data and code publicly available does take some additional time as certain steps have to be taken. For open code, this is mainly around cleaning up and commenting code adequately, although this should be good practice anyway. Publishing code might also encourage scientists to improve their coding abilities (Easterbrook 2014). Data need to be anonymized prior to publication. However, many data sets are already collected in an anonymous way e.g. survey data using an online platform need to be anonymized to comply with GDPR, hence the additional time commitment is often not that large. For personal data, such as smart meter data, a solution might be to only allow access via a secure research portal. A particular challenge for energy research is that projects are often done with an external partner, e.g. a utility company who might oppose publication of data sets for competitive reasons. In the ideal case, this would be addressed prior to commencing any research and an agreement reached on which data (if not the total data set) could be shared, e.g. (Nicolson et al. 2017). For video and audio recordings, data need to be transcribed, removing all personally identifiable features, and / or the sound of the audio-source modified to avoid recognizability (Pätzold 2005). Again, this takes resources but might be required even for institutional storage. We suggest that (where applicable) exact copies of survey and interview questions should be published alongside code and data.

The webpage [re3data.org](https://re3data.org) provides a list of more than 2000 research data repositories that cover research data repositories from different academic disciplines; however, this includes institutional webpages which cannot be accessed by everyone. It is also possible to publish data as a manuscript such as in Nature Scientific Data (see e.g. for

an example of a special collection on energy related data Huebner and Mahdavi, 2019). Similarly, a range of platforms exist for hosting and sharing code, such as [github](https://github.com/) (<https://github.com/>). Some repositories are suitable for a range of uses under one project, such as the Open Science Framework which allows preregistration of studies, depositing of data and code. OpenEI is a data repository specifically for energy related data sets; whilst providing easy access, it does not generate a DOI.

When preparing data and code for sharing, the confidentiality of research participants must be safeguarded, i.e. all data need to be de-identified (for further details, see UK Data Service n.d) .

The FAIR principles state that data should be Findable, Accessible, Interoperable, and Reusable (Wilkinson et al. 2016).

- **Findable.** Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.
- **Accessible.** Once the user finds the required dataset, they must be able to actually access it, possibly including authentication and authorisation.
- **Interoperable.** The data usually need to be integrated with other data. In addition, the data need to interoperate with as wide as possible a variety of applications or workflows for analysis, storage, and processing.
- **Reusable.** To be able to reuse data, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

In addition, a meta document, i.e. a 'ReadMe' document needs to be prepared and / or code documentation, and uploaded together with the data.

### 3. Practical suggestions on tool implementation.

In summary, we suggest four tools to be used by energy researchers to create transparent, reproducible, and quality research. Figure 2 shows where in the research process the tools should be deployed and how they help to overcome problems within the scientific process.

## Improving research practices in energy (preprint 20 July 2020)

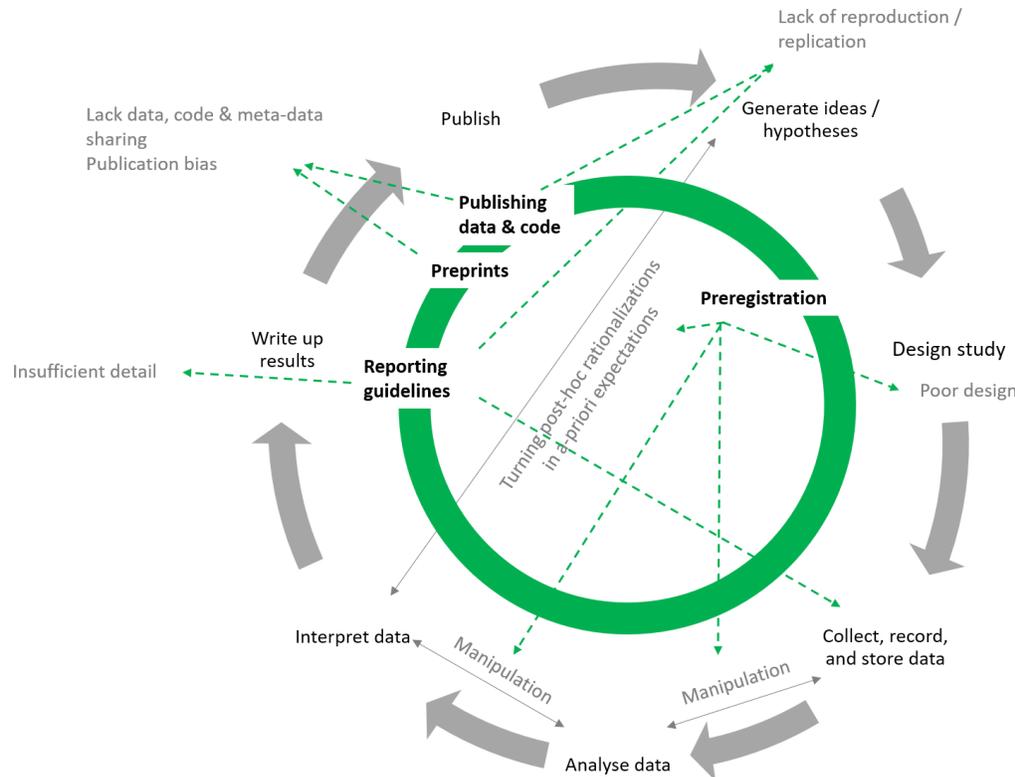


Figure 2. Temporal implementation of the suggested tools and how they help to overcome problematic research practices.

Preregistration happens early in the research process but has impacts for many subsequent stages. It helps to identify problems in the study design when planning how to conduct and analyse the study and can hence improve *poor design*. It also mitigates issues around *manipulation* of data. Through stating what outcomes are expected, preregistration reduces the likelihood of *turning post-hoc rationalizations into a-priori expectations*.

Reporting guidelines are used when results are written up and reduce the problem of *insufficient reporting* of study details. When considered early on, they help to *collect, record, and store data* in a suitable manner. They also make it easier to conduct *reproduction / replication* studies by providing necessary information.

## Improving research practices in energy (preprint 20 July 2020)

Preprints and data and code sharing happen towards the end of the research process. They can happen before publication or together with publication, and in some cases even after publication (if e.g. keeping an embargo on data). Preprints contribute to mitigating the *publication bias*, i.e. negative findings which tend to be published in journals to a lesser extent, can always find a home in preprints and hence a citable, traceable, and indexed record of a study irrespective of outcome can be created. Data, code and meta-data sharing helps to overcome issues around the *lack of sharing* of these items and allow results to be checked, expanded, and synthesized. They make it easier to run *reproduction / replication studies* by providing crucial information for those.

We have developed a simple checklist to suggest how researchers could report on the tools employed in their research in academic publications (see Table 2 and supplementary material).

Table 2. The TReQ list – a checklist for reporting of tools employed to promote transparency, reproducibility, and quality of research.

<b>Tools</b>	<b>Check</b>	<b>Comments</b>
<b><i>Pre-registration</i></b>		
This study has pre-analysis plan.	Yes No	<i>Explanation if no:</i>
<i>If yes</i>		
URL		<i>{insert link}</i>
Was it registered before data collection?	Yes No	<i>Explanation if no:</i>
Does the paper mention and explain deviations from the PAP?	Yes No Not applicable (no deviations)	<i>If yes, specify section of paper, or explanation if no:</i>
<b><i>Reporting guidelines</i></b>		
This paper follows a reporting guideline.	Yes No	
<i>If yes</i>		
Which one?		<i>{insert name and citation, include in reference list}</i>
<b><i>Open Data and Code</i></b>		

## Improving research practices in energy (preprint 20 July 2020)

Data/code are publicly available	Yes, data and code Yes, data only Yes, code only No	
Does the paper make a statement on data and code availability?	Yes, on data and code Yes, on data only Yes, on code only No	<i>Refer to relevant section or include here:</i>
<i>If yes</i>		
What is / are the link(s)?		<i>{insert link(s)}</i>
Have steps been taken to ensure the data are FAIR?	Yes No	
Has meta-data been uploaded?	Yes No	
<b>Preprints</b>		
Have you uploaded a preprint?	Yes Planned following submission No	
<i>If yes</i>		
What is the link?		<i>{insert link}</i>
<i>If planned</i>		
Which preprint server/location?		

We suggest that authors attach a filled template to their paper as an Appendix. It will a) aid others to make a quick judgement on to what extent the work used good practices related to transparency, reproducibility, and rigour, b) aid others to find all additional resources easily, and c) help author(s) to check to what extent they have followed good research practices and what additional actions they could take.

## 4. Discussion and conclusion

In this paper, we suggest four easy-to-use, widely applicable tools to improve the transparency, quality, and reproducibility of research. We provide information and guidance on pre-analysis plans, reporting guidelines, open data and code, and preprints, and have developed a checklist researchers could use to show how they have made use of these tools. We believe that almost all energy researchers could easily (and should) now integrate all or most of these tools, or at least consider and justify not doing so. As we have argued, this should result in a range of benefits for researchers and research users alike.

We recognise that these four only reflect a selection of possible tools. We chose them to be as widely applicable as possible so that researchers using different methods could apply them. Other approaches are also worthy of wider uptake. For example, a little-discussed but important issue concerns the quality of literature reviews, both when undertaken as part of the foundation of any normal research project, or as a method in its own right. Missing or misrepresenting work again risks needless duplication of research, or promulgation of wrong interpretations of previous work (Phelps 2018). A solution is to employ more systematic approaches, which employ structured methods to minimise bias and omissions (Grant and Booth 2009). However, we do not cover systematic review here because the main TReQ aspects of such reviews (such as reporting of search strategies, inclusion criteria, etc.) are not usually found in the background literature review sections of standard articles -- rather they are restricted to review articles where the review is the research method. Similarly, some researchers, such those working with large quantitative data set, might consider simulating data on which to set up and test an analysis before deploying it on the real data.

We also stress that there a number of other practices that whilst not necessarily directly contributing to TReQ are crucial for good research. These include ethics, accurate acknowledgements of the contribution of other researchers, being open about conflict of interests, and following legal obligations of employers. Most universities provide details and policies on academic conduct that is expected from all staff (National Academy of

Sciences 1992). Crucially, while we argue that employing all these approaches (including TReQ) are now *necessary conditions* for research quality, they are not on their own *sufficient conditions*. For example, how findings from a study translate to other times, situations, and people, i.e. its external validity, will strongly depend on study design and method.

Another important point to consider is how to ensure that the tools are not simply deployed as ‘window dressing’, and misused to disguise ongoing questionable research practices under a veneer of transparency. For example, an analysis of pre-analysis plans showed that a significant proportion failed to specify necessary information e.g. on covariates or outlier correction, and that corresponding manuscripts did not report all prespecified analyses (Ofosu and Posner 2019).

To ensure that both open data and open code are useable for others, detailed documentation needs to be provided. Standards for meta-documentation should be considered (e.g. Digital Curation Centre n.d, Centre for Government Excellence n.d). It has been shown that poor quality data is prevalent in large databases and on the Web (Saha and Srivastava, 2014). Regarding reporting guidelines, a potential issue arises from their relatively unfamiliarity both to authors and reviewers. Hence, there is the risk that they are only partially or poorly implemented and that reviewers will have difficulty evaluating them. Even within the medical field, in which reporting guidelines are widely used, research showed that only half of reviewer instructions mentioned reporting guidelines and not necessarily in great detail (Hirst and Altman 2012).

Given that the focus on this paper is to present a general introduction applicable to a wide range of research methods, detailed specifications for specific approaches are beyond its scope. We suggest that especially those who conduct quantitative, experimental research follow detailed suggestions provided elsewhere (e.g. McKenzie 2012; Chuang and Wykstra 2015).

Finally, while we attempt in this paper to convince individual readers of the benefits of adopting these tools, we acknowledge that widespread use is unlikely to come about without more structural adjustments to the energy research ecosystem. This is needed in a number of areas. Tools for better research practices should become a standard part

of (at least) postgraduate education programmes and be encouraged by thesis supervisors. Journals, funders, academic and other research institutions, and regulatory bodies have an important role to play in increasing knowledge and usage. This could involve approaches like signposting reporting guidelines (Simera et al. 2010) and working with editors and peer reviewers to increase awareness of the need for such approaches. Further examples include partnerships between funders and journals, with funders offering resources to carry out research accepted by the journal as a registered report (see Nosek 2020).

In summary, energy research is a field in which it is of utmost importance that research findings can be trusted, given the field's relevance to mitigating and adapting to climate change. At the time of writing, energy research is lagging behind other disciplines when it comes to research transparency, reproducibility, and quality. We present four tools to help improve energy research that are easy to use and can be adapted to a variety of research methods: pre-registration of studies, open data and open code, reporting guidelines, and preprints. We set out briefly how they are used, and why their use is beneficial both for researchers and research users. We have developed a template for researchers on which they could indicate how they have used tools. We call for researchers, and the broader energy research ecosystem, to now accept the use of these and related tools as part of good research practice.

### **Acknowledgements**

GMH and MF gratefully acknowledge support from UK Research and Innovation through the Centre for Research into Energy Demand Solutions, grant reference number EP/R035288/1.

NW is supported by London-Loughborough (LoLo) EPSRC Centre for Doctoral Training in Energy Demand EP/L01517X/1.

### **Authors' contributions**

GMH: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration.

Improving research practices in energy (preprint 20 July 2020)

MF: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Visualization.

NW: Writing – Original Draft, Writing - Review & Editing.

### Data availability

No data was used in this paper.

### References

Baker, M. and Penny, D. (2016) 'Is there a reproducibility crisis?', *Nature*. Nature Publishing Group, pp. 452–454. doi: 10.1038/533452A.

Ball, C. (2009) 'What Is Transparency?', *Public Integrity*. Taylor and Francis Ltd., 11(4), pp. 293–308. doi: 10.2753/PIN1099-9922110400.

Banks, G. C., Rogelberg, S. G., Woznyj, H. Landis, R., Rupp. D. (2016) 'Editorial: Evidence on Questionable Research Practices: The Good, the Bad, and the Ugly', *Journal of Business and Psychology*. Springer New York LLC, pp. 323–338. doi: 10.1007/s10869-016-9456-7.

Banks, G. C., O'Boyle, E. H., Pollack, J. M. (2016) 'Questions About Questionable Research Practices in the Field of Management', *Journal of Management*. SAGE Publications Inc., 42(1), pp. 5–20. doi: 10.1177/0149206315619011.

Begley, C. G. and Ellis, L. M. (2012) 'Drug development: Raise standards for preclinical cancer research', *Nature*. Nature Publishing Group, 483(7391), pp. 531–533. doi: 10.1038/483531a.

Button, K. S., John, P. A., Ioannidis, C. M., Nosek, B. A., Flint, J., Robinson, E. S. J., Munafò, M. R. (2013) 'Power failure: why small sample size undermines the reliability of neuroscience.', *Nature reviews. Neuroscience*. Nature Publishing Group, 14(5), pp. 365–76.

Center for Government Excellence (n.d.) Open Data Metadata Guide. Retrieved 8 July 2020 from <https://labs.centerforgov.org/guides/>

Chambers, C. D., Feredoes, E., Muthuskumaraswamy, D. S., Etchells, P. (2014)

Improving research practices in energy (preprint 20 July 2020)

'Instead of "playing the game" it is time to change the rules: Registered Reports at AIMS Neuroscience and beyond', *AIMS Neuroscience*, pp. 4–17. doi: 10.3934/Neuroscience.2014.1.4.

Chuang, E. and Wykstra, S. (2015) A Guide to Pre-Analysis Plans. Retrieved July 8 2020 from <https://osf.io/sh4v8/>

Collins, G.S., Reitsma J.B., Altman D.G., Moons K.G., (2015) 'Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement', *BMJ (Online)*. BMJ Publishing Group, 350. doi: 10.1136/bmj.g7594.

Crüwell, S. and Evans, N. J. (2019) 'Preregistration in Complex Contexts: A Preregistration Template for the Application of Cognitive Models.' PsyArXiv. doi: 10.31234/OSF.IO.

Dickersin, K., Chan, S., Chalmers, T. C., Sacks, H. S., Smith, H. (1987) 'Publication bias and clinical trials', *Controlled Clinical Trials*. Elsevier, 8(4), pp. 343–353. doi: 10.1016/0197-2456(87)90155-3.

Digital Curation Centre (n.d) List of Metadata Standards. Retrieved 8 July 2020 from <https://www.dcc.ac.uk/guidance/standards/metadata/list>

Easterbrook, S. M. (2014) 'Open code for open science?', *Nature Geoscience*. Nature Publishing Group, pp. 779–781. doi: 10.1038/ngeo2283.

Eisner, D. A. (2018) 'Reproducibility of science: Fraud, impact factors and carelessness', *Journal of Molecular and Cellular Cardiology*. Academic Press, pp. 364–368. doi: 10.1016/j.yjmcc.2017.10.009.

Equator Network (2016) *EQUATOR Reporting Guideline Decision Tree Which guidelines are relevant to my work?* Retrieved July 7<sup>th</sup> 2020 from <https://www.equator-network.org/wp-content/uploads/2013/11/20160226-RG-decision-tree-for-Wizard-CC-BY-26-February-2016.pdf>

Eysenbach, G. (2006) 'The Open Access Advantage' *J Med Internet Res*, 8(2), p.e.8. doi: 10.2196/jmir.8.2.e8

- Fell, M. J. (2019) 'The economic impacts of open science: A rapid evidence assessment', *Publications*. MDPI AG, p. 46. doi: 10.3390/PUBLICATIONS7030046.
- Ferguson, C. J. and Brannick, M. T. (2012) 'Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses', *Psychological Methods*. Psychol Methods, 17(1), pp. 120–128. doi: 10.1037/a0024445.
- Grant, M. J. and Booth, A. (2009) 'A typology of reviews: An analysis of 14 review types and associated methodologies', *Health Information and Libraries Journal*. John Wiley & Sons, Ltd, pp. 91–108. doi: 10.1111/j.1471-1842.2009.00848.x.
- Haven, T. L. and Van Grootel, D. L. (2019) 'Preregistering qualitative research', *Accountability in Research*. Taylor and Francis Inc., 26(3), pp. 229–244. doi: 10.1080/08989621.2019.1580147.
- Hirst, A. and Altman, D. G. (2012) 'Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals', *PLoS ONE*. Public Library of Science. doi: 10.1371/journal.pone.0035621.
- Huebner, G. M., Nicolson, M., Fell, M., Shipworth, D., Elam, S., Hanmer, C., Kennard, H., Johnson, C. (2017 May) 'Are we heading towards a replicability crisis in energy efficiency research? A toolkit for improving the quality, transparency and replicability of energy efficiency impact evaluations'. Paper presented at ECEEE Summer Study Proceedings, Belambra Presqu'île de Giens, France.
- Huebner, G. M. and Mahdavi, A. (2019) 'A structured open data collection on occupant behaviour in buildings', *Scientific Data*, 6(1), p. 292. doi: 10.1038/s41597-019-0276-2.
- Huston, P., Edge, V. and Bernier, E. (2019) 'Reaping the benefits of Open Data in public health', *Canada Communicable Disease Report*. Infectious Disease and Control Branch (IDPCB) - Public Health Agency of Canada, 45(10), pp. 252–256. doi: 10.14745/ccdr.v45i10a01.
- John, L. K., Loewenstein G. and Prelec D. (2012) 'Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling', *Psychological Science*, 23(5), pp. 524-532. doi: 10.1177/0956797611430953

Improving research practices in energy (preprint 20 July 2020)

- Kerr, N. L. (1998) 'HARKing: Hypothesizing After the Results are Known', *Personality and Social Psychology Review*, 2(3), pp. 196-217. doi: 10.1207/s15327957pspr0203\_4
- Kjærgaard, M. B., Ardakanian, O., Carlucci, S., Dong, B., Firth, S. K., Gao, N., Huebner, G. M., Mahdavi, A., Rahaman, M. S., Salim, F. D., Sangogboye, F. C., Schwee, J. H., Wolosiuk, D., Zhu, Y. (2020) 'Current practices and infrastructure for open data based research on occupant-centric design and operation of buildings', *Building and Environment*. Elsevier Ltd, 177, p. 106848. doi: 10.1016/j.buildenv.2020.106848.
- Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., Glennerster, R., Green, D. P., Humphreys M., Imbens, G., Laitin, D., Madon, T., Nelson, L., Nosek, B. A., Petersen, M., Sedlmayr, R., Simmons, J. P., Simonsohn, U., Van der Laam, M. (2014) 'Promoting Transparency in Social Science Research', *Science*, 343(6166), pp.30-31. doi: 10.1126/science.1245317
- McGrath, C. and Nilsson, G. (2018) 'Data sharing in qualitative research: opportunities and concerns', *MedEdPublish*. Association for Medical Education in Europe (AMEE), 7(4). doi: 10.15694/mep.2018.0000255.1.
- McKenzie, D. (2012) 'Beyond baseline and follow-up: The case for more T in experiments', *Journal of Development Economics*, 99(2), pp. 210-221. doi: 10.1016/j.jdeveco.2012.01.002
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. (2009) 'Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement', *BMJ (Online)*. British Medical Journal Publishing Group, pp. 332–336. doi: 10.1136/bmj.b2535.
- Munafò M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn U., Wagenmakers, E., Ware, J. J., Ioannidis J. P. A. (2017) 'A manifesto for reproducible science', *Nature Human Behaviour*, 1(0021), pp.1-9. doi: 10.1038/s41562-016-0021
- National Academy of Sciences (1992) 6 STEPS TO ENCOURAGE RESPONSIBLE RESEARCH PRACTICES in *Responsible Science: Ensuring the Integrity of the Research Process: Volume I*. Washington D.C.: National Academy Press. Retrieved June 6 2020 from <https://www.nap.edu/read/1864/chapter/8>

Improving research practices in energy (preprint 20 July 2020)

Nicolson, M., Huebner, G. M., Shipworth, D., Elam, S. (2017) 'Tailored emails prompt electric vehicle owners to engage with tariff switching information', *Nature Energy*, 2(6), pp. 1–6. doi: 10.1038/nenergy.2017.73.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., Mellor, D. T. (2018) 'The preregistration revolution', *PNAS*, 115(11), pp. 2600-2606. doi: 10.1073/pnas.1708274114

Office for National Statistics (2020) Gross domestic expenditure on research and development, UK: 2018. Retrieved July 8 2020 from <https://www.ons.gov.uk/economy/governmentpublicsectorandtaxes/researchanddevelopmentexpenditure/bulletins/ukgrossdomesticexpenditureonresearchanddevelopment/2018>

Ofose, G. K. and Posner, D. N. (2019) *Pre-analysis Plans: A Stocktaking* \*.

Olken, B. A. (2015) 'Promises and perils of pre-analysis plans', in *Journal of Economic Perspectives*. American Economic Association, pp. 61–80. doi: 10.1257/jep.29.3.61.

Open Science Collaboration (2015) 'Estimating the reproducibility of psychological science', *Science*. American Association for the Advancement of Science, 349(6251), pp. aac4716–aac4716. doi: 10.1126/science.aac4716.

Pätzold, H. (2005) 'Secondary analysis of audio data. Technical procedures for virtual anonymization and pseudonymization', *Forum Qualitative Sozialforschung*, 6(1). doi: 10.17169/fqs-6.1.512.

Pfenninger, S., DeCarolis, J., Hirth, L., Sylvian, Q., Staffell, I. (2017) 'The importance of open data and software: Is energy research lagging behind?', *Energy Policy*. Elsevier Ltd, 101, pp. 211–215. doi: 10.1016/j.enpol.2016.11.046.

Pfenninger, S., Hirth, L., Schlecht, I., Schmid, E., Wiese, F., Brown, T. Davis, C., Gidden, M., Heinrichs, H., Heuberger, C., Hilpert, S., Krien, U., Matke, C., Nebel, A., Morrison, R., Müller, B., Pleßmann, G., Reeg, M., Richstein, J. C., ... Wingenbach, C. (2018) 'Opening the black box of energy modelling: Strategies and lessons learned', *Energy Strategy Reviews*, 19, pp. 63–71. doi: 10.1016/j.esr.2017.12.002.

Phelps, R. P. (2018) To save the research literature, get rid of the literature review. LSE Impact Blog. Retrieved July 8 2020 from

<https://blogs.lse.ac.uk/impactofsocialsciences/2018/06/12/to-save-the-research-literature-get-rid-of-the-literature-review/>

Pienta, A. M., Lyle, J., Alter, G. (2010 April) The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data. Paper presented at The Organisation, Economics and Policy of Scientific Research, Torino, Italy.

Piwohar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., Haustein, S. (2018) 'The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles', *PeerJ*. PeerJ Inc., 2018(2). doi: 10.7717/peerj.4375.

Piwohar, H. A. and Vision, T. J. (2013) 'Data reuse and the open data citation advantage', *PeerJ*. PeerJ Inc., 2013(1), p. e175. doi: 10.7717/peerj.175.

Prinz, F., Schlange, T. and Asadullah, K. (2011) 'Believe it or not: How much can we rely on published data on potential drug targets?', *Nature Reviews Drug Discovery*. Nature Publishing Group, pp. 712–713. doi: 10.1038/nrd3439-c1.

Nosek, B. A., (2020 January) *Changing a Research Culture Toward Openness and Reproducibility* [Presentation at University College London]. London. <https://osf.io/rwn4a/>

Rosenthal, R. (1979) 'The file drawer problem and tolerance for null results', *Psychological Bulletin*, 86(3), pp. 638–641. doi: 10.1037/0033-2909.86.3.638.

Saha, B. and Srivastava, D. (2014) 'Data quality: The other face of Big Data', in *Proceedings - International Conference on Data Engineering*. IEEE Computer Society, pp. 1294–1297. doi: 10.1109/ICDE.2014.6816764.

Schloss, P. D. (2018) 'identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research', *mBio*. American Society for Microbiology, 9(3). doi: 10.1128/mBio.00525-18.

Schulz, K. F., Altman, D. G. and Moher, D. (2010) 'CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials', *BMJ (Online)*. BMJ, 340(7748), pp. 698–702. doi: 10.1136/bmj.c332.

## Improving research practices in energy (preprint 20 July 2020)

- Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K. F., Altman, D. G. (2010) 'Transparent and accurate reporting increases reliability, utility, and impact of your research: Reporting guidelines and the EQUATOR Network', *BMC Medicine*. BioMed Central, p. 24. doi: 10.1186/1741-7015-8-24.
- Simmons J. P., Nelson, L. D., Simonsohn, U. (2011) 'False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant', *Psychological Science*, 22(11), pp.1359-1366. doi: 10.1177/0956797611417632
- Stroebe, W., Postmes, T. and Spears, R. (2012) 'Scientific Misconduct and the Myth of Self-Correction in Science.', *Perspectives on psychological science : a journal of the Association for Psychological Science*. SAGE PublicationsSage CA: Los Angeles, CA, 7(6), pp. 670–88. doi: 10.1177/1745691612460687.
- Suber, P. (2013) *Open Access*. Cambridge: MIT Press. Retrieved July 8 2020 from <https://mitpress.mit.edu/books/open-access>
- Tennant, J. P., Waldner, F., Jacques, D., C., Masuzzo, P., Collster, L. B., Hartgerink, C. H. J. (2016) 'The academic, economic and societal impacts of Open Access: an evidence-based review', *F1000Research*. F1000 Research, Ltd., 5, p. 632. doi: 10.12688/f1000research.8460.3.
- Tong, A., Sainsbury, P. and Craig, J. (2007) 'Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups', *International Journal for Quality in Health Care*, 19(6). doi: 10.1093/intqhc/mzm042.
- UK Data Service (n.d). Anonymisation. Retrieved June 6 2020 from <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation.aspx>
- UK Research and Innovation (n.d) Policy on access to research publications. Retrieved June 6 2020 from <https://www.epsrc.ac.uk/about/access/roaccess/>
- Vandewalle, P. (2012) 'Code sharing is associated with research impact in image processing', *Computing in Science and Engineering*, pp. 42–47. doi: 10.1109/MCSE.2012.63.
- Wagenmakers, E.-J. and Dutilh, G. (2016) 'Seven Selfish Reasons for Preregistration',

*APS Observer*, 29(9), pp. 13–14.

Warren, M. (2018) 'First analysis of "pre-registered" studies shows sharp rise in null findings', *Nature*. Springer Science and Business Media LLC. doi: 10.1038/d41586-018-07118-1.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016) 'Comment: The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*. Nature Publishing Groups, 3. doi: 10.1038/sdata.2016.18.