

## **Evaluation of Indicators of Reproducibility and Transparency in Published Cardiology**

### **Literature**

**Short Title:** Anderson et al.; Indicators of Reproducibility in Cardiology

J. Michael Anderson, B.S.<sup>1</sup>, Bryan Wright, B.G.S.<sup>1</sup>, Daniel Tritz, B.S.<sup>1</sup>, Jarryd Horn, B.S.<sup>1</sup>, Ian Parker, D.O.<sup>2</sup>, Daniel Bergeron, D.O.<sup>2</sup>, Sharolyn Cook, D.O.<sup>2</sup>, Matt Vassar, PhD<sup>1</sup>

1. Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma, USA
2. Oklahoma State University Medical Center - Department of Cardiology, Tulsa, Oklahoma, USA

**Corresponding Author:** Mr. J. Michael Anderson, Oklahoma State University Center for

Health Sciences, 1111 W 17<sup>th</sup> St Tulsa, OK 74107, United States;

Fax: 918-561-8428; Telephone: 918-521-8774; Email: [jande31@okstate.edu](mailto:jande31@okstate.edu)

Word Count: 7041

## **Abstract:**

### **Background**

The extent of reproducibility in cardiology research remains unclear. Therefore, our main objective was to determine the quality of research published in cardiology journals using eight indicators of reproducibility.

### **Methods**

Using a cross-sectional study design, we conducted an advanced search of the National Library of Medicine (NLM) catalog for publications from 2014-2018 in journals pertaining to cardiology. Journals must have been published in the English language and must have been indexed in MEDLINE. Once the initial list of publications from all cardiology journals was obtained, we searched for full-text PDF versions using Open Access, Google Scholar, and PubMed. Studies were analyzed using a pilot-tested Google Form to evaluate the presence of information that was deemed necessary to reproduce the study in its entirety.

### **Results**

After exclusions, we included 132 studies containing empirical data. Of these studies, the majority (126/132, 95.5%) did not provide the raw data collected while conducting the study, 0/132 (0%) provided step-by-step analysis scripts, and 117/132 (88.6%) failed to provide sufficient materials needed to reproduce the study.

### **Conclusions**

The presentation of studies published in cardiology journals does not appear to facilitate reproducible research. Considerable improvements to the framework of biomedical science, specifically in the field of cardiology, are necessary. Solutions to increase the reproducibility and

transparency of published works in cardiology journals is warranted, including addressing inadequate sharing of materials, raw data, and key methodological details.

## **Introduction:**

Cardiology is a field at the forefront of evidence-based medicine. Advances in diagnostic and therapeutic modalities related to imaging, biomarkers, drugs, devices, and minimally invasive procedures have improved clinical outcomes and quality of life.<sup>1</sup> However, disagreements exist between the results from randomized clinical trials and observational studies – two study designs that often provide the evidentiary base for clinical practice guidelines in cardiology. In such cases, some authors have argued that “evidentiary policies should emphasize the open access to clinical data [and] the responsible analysis of these data by agencies immune to the influence of partisan stakeholders”<sup>2</sup> to better account for disparate study findings and to improve reproducibility. Take as an example the United States Preventive Services Task Force’s (USPSTF) and American Heart Association’s (AHA) clinical practice recommendations for aspirin prophylaxis to prevent cardiovascular disease. The original recommendations were based on a meta-analysis of five studies that found a 28% decrease in coronary disease, with no improvement in total mortality, and an increased risk of bleeding.<sup>3,4</sup> A more recent meta-analysis of 11 studies of aspirin for primary prevention found no reduction in mortality and a 0.6% increase in major bleeding.<sup>5</sup> This most recent analysis lead to a reversal in the clinical recommendations for patients >70 years old and called into question its use for patients between 40 and 70 years of age.<sup>6</sup> One possible mechanism to prevent similar reversals of practice in cardiovascular medicine is to advocate for reproducible, transparent research practices.

Concern over the lack of reproducibility in scientific research has been well documented.<sup>7-10</sup> Attention surrounding this issue has grown to such prominence that a recent survey in *Nature* reported that 90% of respondents from the scientific community agreed that a ‘reproducibility crisis’ exists.<sup>11</sup> The National Institutes of Health (NIH) recently called for “immediate and substantive action” to address this shortcoming.<sup>12</sup> This call to action stems from reports that only 10–25% of scientific experiments are reproducible,<sup>8,11,13</sup> resulting in \$28 billion per year spent on irreproducible preclinical research in the United States alone.<sup>14</sup> In 2018, the National Academies of Science, Engineering, and Medicine drafted a congressionally-mandated report, *Reproducibility and Replicability in Science*<sup>15</sup>, which outlines recommendations for research stakeholders – including researchers, journals, funding agencies, and academic institutions – to improve the reproducibility and transparency of scientific research. Prior to release of the National Academies’ report, this body, referred to as the Institute of Medicine at the time, established the Committee on Strategies for Responsible Sharing of Clinical Trial Data that recommended all studies with a novel study design provide detailed data sharing, metadata, protocols, statistical analysis plans, and analytic code to be considered reproducible.<sup>16,17</sup>

Within the cardiology specialty, some journals have added requirements for authors to share data and append supplementary materials to published reports in an attempt to improve reproducibility. For instance, the American Heart Association journals all adhere to Transparency and Openness Promotion (TOP) Guidelines that outline mechanisms for sharing data, statistical code, and other materials to facilitate reproducibility.<sup>18</sup> The urgency for reproducible research highlights the need for closer examination of practices within the field of cardiology. In this study, we evaluate eight indicators of reproducibility and transparency posited by Hardwicke et al.<sup>19</sup> in a random sample of studies published in cardiology journals.

Results from this investigation showcase areas where improvements are most needed and serve as baseline data for future, comparative investigations.

## **Methods:**

This investigation employed an observational, cross-sectional design. Since the Open Science Framework has generated a new infrastructure that advocates for research transparency, we have supplied our complete protocol, raw data, and other necessary materials at <https://osf.io/x24n3/>. Using a similar methodology and protocol to that outlined by Hardwicke et al.<sup>19</sup>, with slight modifications, our study focused, specifically, on the field of cardiology. This study was not subject to institutional review board oversight according to the U.S. Department of Health and Human Services' Code of Federal Regulation 45 CFR 46.102(d) and (f)<sup>20</sup> because our study lacked human subject participation. The present investigation was reported using the guidance for conducting meta-research outlined by Murad and Wang.<sup>21</sup> In addition, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were used when appropriate. Our primary endpoint was to assess for the presence of eight indicators that support transparent research practices, allowing for feasible reproduction of cardiology studies.

### ***Journal and Study Selection***

For this project, one of us (DT) searched all journals in the National Library of Medicine (NLM) catalog using the subject term "Cardiology[ST]". Our search was conducted on May 29, 2019. To meet inclusion criteria, the journals must be in English and MEDLINE indexed. The list of journals generated from our search of the NLM catalog was then extracted by electronic ISSN number (or the linking ISSN if the electronic version was not available). The journal search string of ISSN numbers was then used in PubMed (which encompasses the entire

MEDLINE database) on May 29th, 2019 to gather all publications. For our final list of publications, we included only those from January 1, 2014 to December 31, 2018. Using Excel's random number function, a random number was assigned to all studies. A random sample of 300 studies were chosen for our analysis. This list of studies is available on the Open Science Framework (<https://osf.io/4mtwq/>). To ensure a wide range of publications, we did not restrict any particular study designs from inclusion.

### ***Extraction Training***

Prior to data extraction, two investigators (MA, BW) completed thorough training to establish consistent study evaluation methods between investigators. Training consisted of an in-person session that reviewed study objectives, design, protocol, the Google form used for data extraction, and how to obtain the pertinent information from sampled publications. The methodology was demonstrated by DT using two example studies. MA and BW then independently extracted data from three sample studies in a blinded, duplicate manner. Following this mock data extraction, the pair reconciled all differences between them. Applying the same blinded, duplicate data extraction process, MA and BW extracted data from the first 10 publications of the list generated using the NLM catalog. This was followed by a final consensus meeting to amend all discrepancies. This initial training session was recorded and available online for reference (<https://osf.io/tf7nw/>).

### ***Data Extraction***

Upon completing training, these investigators extracted data from the remaining 290 publications in a blinded, duplicate fashion. Data extraction began on June 3, 2019 and was completed on June 10, 2019. Once data extraction was complete, a final consensus meeting was held by MA and BW to resolve any discrepancies. If necessary, DT was available for

adjudication. We divided publications into two separate categories: (1) publications consisting of empirical data (e.g. clinical trial, cohort, case series, case reports, case-control, secondary analysis, chart review, commentary [with data analysis], and cross-sectional) and (2) studies lacking empirical data (e.g., editorials, commentaries [without reanalysis], simulations, news, reviews, and poems). Using a pilot-tested Google form similar to that created by Hardwicke et al.<sup>19</sup>, with additions, investigators were prompted to identify whether a study had important indicators considered necessary to reproduce a study (<https://osf.io/3nfa5/>). The Google form used to extract data prompted investigators to extract appropriate information based on the individual study design of each publication. Studies lacking empirical data were only analyzed for the presence of funding source(s), disclosure of conflicts of interest, open access, and impact factor (2017, 2018 and five-year impact factor). Case reports and case series were extracted using the same methods as studies that lacked empirical data. Case reports and case series typically lack pre-specified protocols and are, therefore, inherently difficult to reproduce.<sup>22</sup> For systematic reviews/meta-analyses, all indicators were evaluated with the exception of materials availability since it is not expected for them to include additional materials. Our form also expanded the extraction constructed by Hardwick et al. to include the following options for study design: case series, cohort, secondary analysis, systematic review/meta-analysis, chart review, and cross-sectional studies. Finally, the options for funding source were expanded to include more specific identifiers of funding sources that included: university, hospital, public, private/industry, and/or non-profit.

### ***Assessment of Open Access Data***

To assess the public's ability to access each publication included in our sample, we attempted to obtain the full-text PDF version of each publication via a systematic process. First,

we searched for publication title, DOI, and/or PubMed ID using Open Access Button (<https://openaccessbutton.org/>). If this search method was unsuccessful, investigators then searched Google Scholar and PubMed for publication title, DOI, and PubMed ID. If the full-text was still unobtainable after all search methods were exhausted, the publication was considered to be paywall restricted and, therefore, not accessible to the public.

### ***Evaluation of Replication and Whether Publications Were Included in Research Synthesis***

For studies with empirical data, we searched Web of Science to determine whether the present study: (1) was included as part of a replication study in subsequent publications and (2) had been cited in systematic reviews and/or meta-analyses. By reviewing the titles and abstracts of all publications in which the study was cited, we made the determination of the frequency with which the given study was cited in replication studies or systematic reviews/meta-analyses. This process was performed using a blinded, duplicate data extraction process.

### ***Analysis Plan***

Outcomes are presented as percentages with 95% confidence intervals. We reported descriptive statistics for each category using Microsoft Excel.

### **Results:**

Our NLM search identified 280 journals for cardiology. Only 150 journals met the inclusion criteria and yielded 145 471 publications from 2014 to 2018. A random sample of 300 publications yielded 290 eligible publications and 10 inaccessible ones (Figure 1). The variables analyzed depended on individual study designs. Thus, Supplementary Table 1 outlines a detailed description of the included publications and the corresponding analysis of the eight queried indicators of reproducibility.

### ***Sample Characteristics***

Our final analysis included publications lacking empirical data ( $n = 128$ ), including editorials, commentaries (without reanalysis), simulations, news, and reviews. Even though case reports and case series ( $n = 30$ ) provide empirical data, the given methodologies are typically either insufficient for reproduction or absent (as detailed by Wallach et al.<sup>22</sup>). We, therefore, extracted these study designs in an identical fashion as studies with non-empirical data. The common absence of reproducible methodology in case reports and case series formulated the basis for excluding these study types from certain analyses. Thus, our final sample size consisted of 132 publications with empirical data from reproducible study types (e.g. cost-effectiveness/decision-making, clinical trial, cohort, case-control, secondary analysis, chart review, commentary [with data analysis], survey, laboratory, and cross-sectional). Studies lacking empirical data made up the largest percentage (44.1%, 128/290) of studies included in our sample. The median five-year journal impact factor was 4.023 (range: 0.05-23.425). Journal impact factors were inaccessible for 16 publications (Supplemental Table 2). Additional reproducibility and transparency characteristics for our sample of publications from cardiology journals are available in Table 1 and Supplemental Table 3.

### ***Article Availability***

Among the 290 accessible publications, 168 (58%) were publicly available, whereas 122 (42%) were only available through a paywall (Supplemental Table 3). Furthermore, we considered the 10 publications for which full-text was unavailable to be paywall restricted. Thus, 132 publications (of 300, 44%) were considered to be restricted behind a paywall.

### ***Materials Availability***

We considered “materials” and “protocol” as any study item and/or document that might be deemed necessary to accurately reproduce the study. Our analysis revealed only 15 publications (of 132) containing empirical data that provided a materials availability statement, and only one publication (of 132) that explicitly stated no supplemental materials were available (Supplemental Table 3). Author statements declaring how to obtain supplemental materials included online journal supplemental appendices (n = 10), an online third party (n = 2), requesting from the corresponding author (n = 2), and a personal or institute website (n = 1). Supplemental materials that authors claimed to be available were only accessible for only seven (47%) of the 15 publications. We did not attempt to verify the availability of materials by requesting them from the corresponding author.

### ***Protocol Availability***

In addition, only four publications (of 127) with empirical data – excluding case reports, case series, and meta-analyses, as detailed in the methods section – provided links to accessible protocols. Of the four publications claiming to provide access to full protocols, two of the links did not successfully navigate to an accessible protocol. The remaining two publications linked to accessible methods and analysis plans (Supplemental Table 3).

### ***Data Availability***

Complete and accurate validation of analyses performed within scientific literature requires access to the raw data collected by a particular study. For the purpose of our study, “raw data” refers to a detailed record of all collected information in an unmodified, digital form (e.g., sampling units of testing participants). Ninety-nine publications containing empirical data lacked data availability statements with the remaining 33 publications containing a data availability statement. Of the 33 publications with data availability statements, 28 of the

referenced supplementary materials were available via the associated online journal, three upon request from the corresponding author(s), and one via an online third party repository (Supplementary Table 3). During data extraction, we were only able to access, download, and open supplementary data files for 19 of the 33 publications claiming additional data availability. Furthermore, only six of the 19 publications with accessible supplementary data provided all raw data necessary to reproduce all calculations and statistical analyses performed by the researchers in the given study. From our sample, less than 4% (6/132) of publications in cardiology journals provide access to complete records of raw numerical data. We did not attempt to verify the 3 publications which claimed to provide data upon request from the corresponding author.

### ***Analysis Script Availability***

Our definition of analysis script encompassed a detailed description of data preparation and either step-by-step instructions for using point--and--click software, analysis code (e.g., Python, Matlab), or syntax (e.g., SPSS, Stata). Our analysis failed to identify a single publication (0 of 132) within our sample that included an analysis script statement.

### ***Preregistration***

Open access registries, such as ClinicalTrials.gov (a public database created by the Food and Drug Administration [FDA] and the National Institutes of Health [NIH]), provide specific information pertaining to a given study, including hypotheses, methods, and/or analysis plans.<sup>23</sup> Less than 5% of publications (6 of 132) with empirical data provided information regarding registration prior to the commencement of the study. In addition, only four of the six provided sufficient information to completely access the study registration. Three preregistered studies (of four, 75%) provided study methods, two (of four, 50%) provided hypotheses, and two (of four, 50%) provided analysis plans (Supplemental Table 1).

### ***Conflict of Interest and Funding Statements***

Among the 290 publications, 241 (83%) provided a conflict of interest (COI) statement. Sixty-four publications (22%) stated the authors had no COI, whereas 177 publications (61%) disclosed authors having one or more COI. Forty-nine publications (17%) did not include a conflict of interest statement. A total of 168 publications (58%) did not include a funding statement, and 34 publications (12%) reported the study was conducted without funding. Of the 122 publications reporting a funding source, the majority (n = 71) of the funds were provided by public sources (Table 1).

### ***Replication and Evidence Synthesis***

Findings from previous studies may be included in future systematic reviews and meta-analyses that can be used to help answer focused clinical questions<sup>24</sup>; however, our analysis shows the publications included in our sample were not commonly cited in other studies. A total of 26 publications (out of 132, 19.7%) with empirical data were cited in one or more subsequent meta-analysis or systematic review (range: 1 to 7; Supplemental Table 4). Our sample did not include a single self-identified replication study.

### **Discussion:**

The climate of studies published in cardiology journals does not appear to facilitate reproducible research. We found that only 25% of publications included data availability statements and only 11% included material availability statements. Despite some publications providing additional data and materials, less than 5% provided the complete raw data used to calculate study findings. Furthermore, we found that publications in cardiology journals do not consistently provide analysis scripts and step-by-step protocols to thoroughly reproduce a study.

Variable analytical approaches to the same data set makes it difficult to reach the same or similar outcomes.<sup>25</sup> Incomplete analysis scripts allow replication studies to employ different analytical approaches with the same data set, which may result in different outcomes.<sup>26</sup> This lack of transparency makes it difficult to verify the accuracy and legitimacy of results. Our findings are similar to a previous study examining reproducibility in biomedical research that concluded certain indicators of reproducibility – including data sharing, materials availability statements, and disclosures of funding sources – are inconsistently reported.<sup>22</sup> The reproducibility surrounding biomedical research will continue to be suboptimal if open access to published literature remains restricted, and the necessary components of a study are not provided, such as all supplemental materials, raw data, and key methodological information.

The majority of studies included in our analysis failed to provide access to all raw numerical data, suggesting a lack of data sharing in cardiology studies, a well-documented obstacle to reproducible research. For example, a survey in *Nature* addressing inaccessible raw data from the original laboratory as a barrier to reproducibility reported that nearly 70% of scientists were unable to successfully reproduce study outcomes in the medical literature.<sup>11</sup> The idea of inferential reproducibility (e.g. rerunning the analyses using the same raw data from the original study), focuses on an essential aspect of research reproducibility. This idea is based upon the notion that an individual should be able to start with the same raw data and end with similar conclusions; however, it has been previously demonstrated that different conclusions may be reached following analysis of the same raw data sets.<sup>27,28</sup> For example, a 2018 study concluded a “significant variation in the results of analyses of complex data may be difficult to avoid, even by experts with honest intentions.”<sup>25</sup> The practice of data sharing has the potential to increase

reproducibility and transparency in biomedical research; however, providing sufficient protocols and analysis scripts used in data analysis is equally important.

Failure to share sufficient protocols and analysis scripts to rerun computations contributes to the non-reproducibility and lack of transparency in research. For example, a 2018 study<sup>28</sup> demonstrated that researchers often repeat a previous experiment using the same methodological steps and protocols to determine the level of reproducibility. . Inadequate methodological reporting and absent analysis scripts makes it difficult to reproduce study outcomes and highlights claims made in previous studies regarding the poor reproducibility of research.<sup>9,10,13</sup> However, it should be noted that another barrier to reproducing scientific literature is a result of digital information becoming “unusable, inoperative, or unavailable because of technological breakdown and evolution or lack of continued curation.”<sup>29</sup> Given that the information necessary to reproduce a study is frequently housed within the published text, the lack of open access to medical literature can also perpetuate non-reproducibility.

Studies available via open access have a greater chance of being subjected to criticism from a broader audience; therefore, lack of open access is a barrier to reproducibility. The inability to view original research materials and protocols in their entirety precludes a comprehensive evaluation of a study.<sup>30</sup> For example, the authors of a 2017 study credit the lack of open access as contributing to the inability to obtain full protocols and analysis scripts.<sup>31</sup> In a sense, open access provides another means of transparency for those who would seek to reproduce a study and validate outcomes. Nearly one-half of the publications included in our sample were inaccessible due to paywall restrictions. This is comparable to a previous study investigating reproducibility in biomedical sciences which found that approximately 43% of publications with a PubMed ID were not publicly available.<sup>22</sup> Preregistration of studies might

help overcome this by making accessible hypotheses, protocols, primary outcome measures.

Preregistering studies on public registries may provide sufficient information to reproduce published and unpublished research alike, even if the full-text is not publicly available.<sup>32</sup>

### ***Future Directions***

The optimum framework by which to address the obstacles surrounding non-reproducibility of studies in cardiology journals is unclear. Requiring authors at the time of publication to submit all raw data and materials necessary to conduct the study might increase the ease with which computations can be reproduced and validated. Efforts have been made to increase the rate at which data is openly shared by authors. For example, the TOP initiative, a 2015 promotion pioneered by Nosek and colleagues<sup>32</sup>, proposes guidelines on data sharing to help increase the dissemination of research data while simultaneously helping increase the reproducibility of scientific literature. Many journals, including Science, implement aspects of the TOP guidelines to hold authors to a higher standard by requiring disclosure of all data used in the analysis such that any researcher may reproduce the findings.<sup>31</sup> Additional journals implementing these same TOP guidelines include the American Heart Association and American Stroke Association journals.<sup>33</sup> More widespread TOP guideline adoption has the potential to minimize the reproducibility crisis. Furthermore, reviewers have taken the initiative to increase the clarity of scientific literature by pledging to adhere to The Peer Reviews' Openness Initiative. By making this pledge, reviewers agree to refuse a comprehensive review of any manuscript from authors who fail to make raw data openly accessible without sufficient reason to the contrary.<sup>34</sup> For these reasons, we urge all journals to consider adopting similar data sharing mandates to make raw data available from all publications, helping to address one cause of non-reproducibility.

While we support initiatives to increase the amount of openly accessible information through data and materials sharing, we also acknowledge the potentially problematic issues that may occur if more raw data is openly available. For example, Lango and Drazen<sup>35</sup> warn that data sharing has the potential for readers to misinterpret or not completely understand the means by which data were collected. Furthermore, they also assert that data sharing allows for the opportunity to drastically increase the number of “Research Parasites” – researchers who had no part in the original collection, analysis, and interpretation of data, but instead use another person’s data for self-gain. Adopting a symbiotic method of data sharing might be a compromise for the research community’s needs. A possible solution has been proposed by Lango and Drazen<sup>35</sup> to permit collaborative data sharing between researchers. This proposal suggests that when a novel idea is posited by a second party researcher, the novel study designer should then locate potential collaborators that have already completed the groundwork data collection. Using the collaborator’s data, both parties work alongside one another on the novel experiment. Not only would this solution ensure accurate representation of previously collected data, but conducting research in this manner may facilitate appropriate co-authorship for all who contributed to the novel investigation. Regardless of the methodology by which data is shared, more transparent and openly accessible information to reproduce any given study is necessary.<sup>35</sup>

In addition, it has been argued that all studies should be preregistered prior to commencement of the study in hopes of increasing transparency in medical research. For example, Munafo et al.<sup>31</sup> suggests the creation of a system whereby the study is preregistered and makes publicly available the following: results, study design, primary outcome, and analysis plan. We agree with the aforementioned proposal, since preregistration might help mitigate analytical manipulation, such as outcome switching and p-hacking – misrepresenting data and/or

statistical analyses in such a way that a non-significant finding is found to be significant.<sup>36</sup> By requiring a more robust standard of reporting analysis plans and step-by-step protocols through preregistration, future researchers might avoid difficulties when trying to replicate a study. Taking steps to address this concern, The Center for Open Science provides badges to publications which are preregistered, include raw data, and all necessary materials.<sup>31</sup> Badges signify the importance these journals place on information availability; however, the influence these badges have on open science practices is not well understood and has been the subject of debate.<sup>37</sup> Despite the unknown influence these badges have on the dissemination of medical research, we encourage a more widespread application of incentive-based guidelines with the goal of increasing the availability of study protocols, materials, and data, even for publications not openly accessible.

### ***Strengths and Limitations***

Our study has many strengths but is not devoid of limitations. The use of open science practices (making available all materials, protocols, analysis plans, and raw data) allows for greater transparency and reproducibility of our study. The blinded, double extraction technique implemented – the gold standard for meta-research data extraction<sup>38</sup> – increases the reliability of our findings. In addition, the robust training provided to investigators prior to initiating this study helped to ensure the reliability of the study results. Despite employing the gold standard to ensure the reliability of data extraction, one limitation of our study includes the possibility that supplemental materials, data, protocols, analysis scripts, and other additional information went unobserved, thereby affecting the outcomes of our study. Second, our analysis of each publication was limited to what was available to investigators through PDF versions and any listed supplemental materials. Additional information necessary to reproduce a study might be

obtainable through correspondence with the primary author; however, we did not attempt to obtain information via such methods. Lastly, the results from this cross-sectional analysis of cardiology publications may not be generalizable to other medical specialties. For these reasons, careful interpretation of our findings as a lower bound estimate of study reproducibility in cardiology journals is warranted.

In conclusion, considerable needs exist regarding the reproducibility of studies in cardiology. Current data sharing, disclosure of materials and analysis scripts, as well as access to step-by-step protocols remain unacceptably low. By failing to make essential aspects of scientific literature available, the reproducibility crisis surrounding biomedical research will continue. To promote further transparency and reproducibility, recommendations are proposed to the editors and to the framework of requirements for article submission in cardiology journals to surmount these shortcomings.

**Sources of Funding:** This study was funded through the 2019 Presidential Research Fellowship Mentor – Mentee Program at Oklahoma State University Center for Health Sciences.

**Disclosures:** The authors report no conflicts of interest.

**Affiliations:** From Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma, USA (J.M.A, B.W., D.T., J.H., M.V.); From Oklahoma State University Medical Center - Department of Cardiology, Tulsa, Oklahoma, USA (I.P., D.B., S.C.)

## References

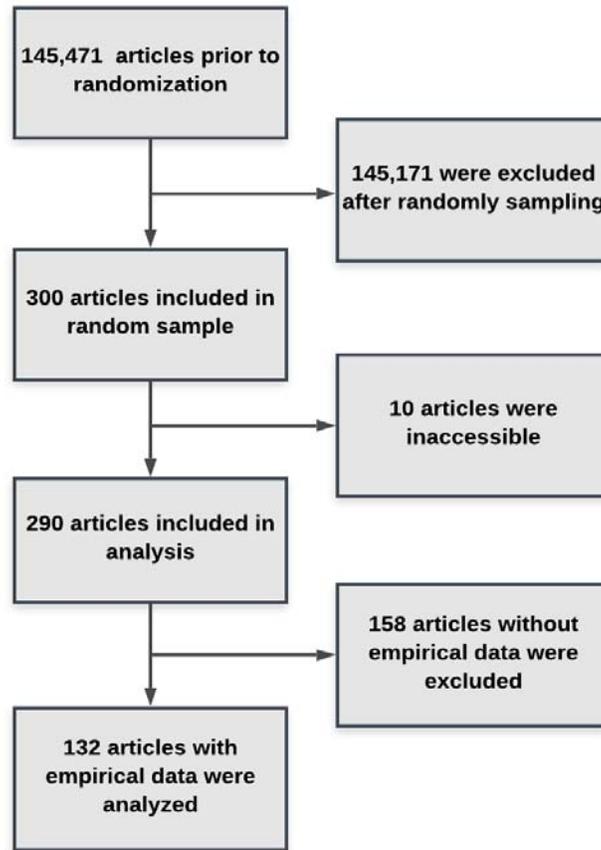
1. Windecker S. Interventional cardiology's golden age of contributions to evidence-based medicine. *EuroIntervention*. 2019;15(1):31-34. doi:10.4244/eijv15i1a6
2. Diamond GA. Randomized trials, observational registries, and the foundations of evidence-based medicine. *Am J Cardiol*. 2014;113(8):1436-1441.
3. U.S. Preventive Services Task Force. Aspirin for the primary prevention of cardiovascular events: recommendation and rationale. *Ann Intern Med*. 2002;136(2):157-160.
4. Pearson TA, Blair SN, Daniels SR, et al. AHA Guidelines for Primary Prevention of Cardiovascular Disease and Stroke: 2002 Update. *Circulation*. July 2002. doi:10.1161/01.CIR.0000020190.45892.75
5. Mahmoud AN, Gad MM, Elgendy AY, Elgendy IY, Bavry AA. Efficacy and safety of aspirin for primary prevention of cardiovascular events: a meta-analysis and trial sequential analysis of randomized controlled trials. *Eur Heart J*. 2019;40(7):607-617.
6. Arnett DK, Blumenthal RS, Albert MA, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease. *Circulation*. March 2019:CIR0000000000000678.
7. McNutt M. Journals unite for reproducibility. *Science*. 2014;346(6210):679-679. doi:10.1126/science.aaa1724
8. Ioannidis JPA. Acknowledging and Overcoming Nonreproducibility in Basic and Preclinical Research. *JAMA*. 2017;317(10):1019-1020.
9. Al-Shahi Salman R, Beller E, Kagan J, et al. Increasing value and reducing waste in

- biomedical research regulation and management. *Lancet*. 2014;383(9912):176-185.
10. Chalmers I, Bracken MB, Djulbegovic B, et al. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014;383(9912):156-165.
  11. Baker M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016;533(7604):452-454.
  12. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature*. 2014;505(7485):612-613.
  13. Begley CG, Ioannidis JPA. Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res*. 2015;116(1):116-126.
  14. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. *PLoS Biol*. 2015;13(6):e1002165.
  15. Science C on RAR in, Committee on Reproducibility and Replicability in Science, Committee on National Statistics, et al. Reproducibility and Replicability in Science. 2019. doi:10.17226/25303
  16. Lo B. Sharing clinical trial data: maximizing benefits, minimizing risk. *JAMA*. 2015;313(8):793-794.
  17. Institute of Medicine, Board on Health Sciences Policy, Committee on Strategies for Responsible Sharing of Clinical Trial Data. *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. National Academies Press; 2015.
  18. Journal Policies | AHA/ASA Journals. <https://www.ahajournals.org/top-guidelines>. Accessed June 22, 2019.

19. Hardwicke TE, Wallach JD, Kidwell M, Ioannidis J. An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014-2017). April 2019. doi:10.31222/osf.io/6uhg5
20. Electronic Code of Federal Regulations- US Department of Health and Human Services' Code of Federal Regulation 45 CFR 46.102(d). [https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML#se45.1.46\\_1102](https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML#se45.1.46_1102) in effect July 19, 2018.
21. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med.* 2017;22(4):139-142.
22. Wallach JD, Boyack KW, Ioannidis JPA. Reproducible research practices, transparency, and open access data in the biomedical literature, 2015–2017. *PLoS Biol.* 2018;16(11):e2006930.
23. Steinbrook R. Public Registration of Clinical Trials. *New England Journal of Medicine.* 2004;351(4):315-317. doi:10.1056/nejmp048191
24. Stevens KR. Systematic reviews: the heart of evidence-based practice. *AACN Clin Issues.* 2001;12(4):529-538.
25. Silberzahn R, Uhlmann EL, Martin DP, et al. Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Advances in Methods and Practices in Psychological Science.* 2018;1(3):337-356.
26. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution.

- Proceedings of the National Academy of Sciences. 2018;115(11):2600-2606.  
doi:10.1073/pnas.1708274114
27. Samsa G, Samsa L. A Guide to Reproducibility in Preclinical Research. *Acad Med*. 2019;94(1):47-52.
  28. Niven DJ, McCormick TJ, Straus SE, et al. Reproducibility of clinical research in critical care: a scoping review. *BMC Med*. 2018;16(1):26.
  29. National Academies of Sciences Engineering, Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press; 2019.
  30. Vazire S. Quality uncertainty erodes trust in science. *Collabra: Psychology*. 2017;3(1).  
<https://collabra.org/articles/10.1525/collabra.74/print/>.
  31. Munafò MR, Nosek BA, Bishop DVM, Button KS, Ioannidis JPA. A manifesto for reproducible science. *Nature Human Behaviour*. 2017;1(1):0021.
  32. Nosek BA, Alter G, Banks GC, et al. SCIENTIFIC STANDARDS. Promoting an open research culture. *Science*. 2015;348(6242):1422-1425.
  33. Journal Policies | AHA/ASA Journals. <https://www.ahajournals.org/TOP-guidelines>.  
Accessed June 25, 2019.
  34. Morey RD, Chambers CD, Etchells PJ. The peer reviewers' openness initiative. 2016.  
<http://lib.pyu.edu.vn/handle/123456789/2221>.
  35. Longo DL, Drazen JM. Data Sharing. *New England Journal of Medicine*. 2016;374(3):276-277. doi:10.1056/nejme1516564

36. Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of p-hacking in science. *PLoS Biol.* 2015;13(3):e1002106.
37. 0000-0002-9089-, 0000-0002-5418-, 0000-0003-1384-, 0000-0002-3125-, 0000-0001-9845-282X, 0000-0001-6797-. Bias in Open Science Advocacy: The Case of Article Badges for Data Sharing | Absolutely Maybe. Absolutely Maybe. <https://blogs.plos.org/absolutely-maybe/2017/08/29/bias-in-open-science-advocacy-the-case-of-article-badges-for-data-sharing/>. Published August 29, 2017. Accessed June 25, 2019.
38. Higgins JPT, Green S, Others. *Cochrane handbook for systematic reviews of interventions*. 2006.



**Figure 1: Flow Diagram for Included and Excluded Studies**

**Table 1. Reproducibility and Transparency Characteristics for Sample of Publications in Cardiology Journals**

Characteristics		Variables
		N (%)
<b>Study Design*</b>	No empirical data (e.g., editorials, commentaries [without reanalysis], simulations, news, and reviews)	128 (44.1)
	Meta-analysis	5 (1.7)
	Commentary with reanalysis	0 (0.0)
	Cost effectiveness and/or decision making	1 (0.3)
	Clinical trial	16 (5.5)
	Case study	30 (10.3)
	Case series	0
	Cohort	46 (16)
	Case control	3 (1.0)
	Survey	1 (0.3)
	Laboratory	26 (9.0)

	Multiple	1 (0.3)
	Other	33 (11.4)
<b>Funding Source*</b>		
	University	1 (0.3)
	Hospital	1 (0.3)
	Public	44 (15.2)
	Private/Industry	10 (3.4)
	Non-profit	1 (0.3)
	No statement listed	169 (58.3)
	No funding received	34 (11.7)
	Mixed	30 (10.3)
<b>Test Subjects*</b>		
	Animals	18 (6.2)
	Humans	114 (39.3)
	Both	0 (0.0)
	Neither	158 (54.5)
<b>Country of Journal Publication*</b>		
	US	177 (61)
	Japan	15 (5.2)
	Germany	6 (2.1)
	France	1 (0.3)

	Italy	2 (0.7)
	Unclear	12 (4.1)
	Other	77 (26.6)
<b>Country of Corresponding Author*</b>		
	US	87 (30)
	China	18 (6.2)
	UK	1 (0.3)
	Germany	15 (5.2)
	Japan	22 (7.6)
	France	7 (2.4)
	Canada	11 (3.8)
	Italy	14 (4.8)
	India	4 (1.4)
	Spain	8 (2.8)
	Unclear	3 (1)
	Other	100 (34.5)
* All empirical and non-empirical studies included in this study (n=290).		

**SUPPLEMENTAL MATERIALS:**

**Supplemental Table 1. Indicators of Reproducibility.**

Analysis of variables within each publication was dependent on study classification. Rationale detailing study classification is available here: <https://osf.io/x24n3/>

<b>Indicators of Reproducibility and Transparency</b>	<b>Study Types Included for Analysis of Reproducibility Indicator</b>	<b>Usefulness for Reproducing Medical Literature</b>
Materials Available	Empirical studies* (n=127)	Providing all materials, including stimuli, survey instruments, and computer code/software used for data collection or running experiments increases the feasibility by which researchers' are able to successfully replicate a study using identical methodology.
Raw Data	Empirical studies† (n=132)	Disseminating records of collected information in its unaltered, digital form facilitates validation of study outcomes through reanalysis

		and helps prevent selective outcome reporting.
Analysis Scripts Available	Empirical studies† (n=132)	Providing access to well-documented, step-by-step instructions detailing how data was prepared and analyzed can increase the clarity of data analysis and reduce inadvertent computations yielding different study outcomes.
Protocol Available	Empirical studies† (n=132)	Access to complete protocols is imperative to efficiently reproduce a study in its entirety. Even the slightest modification to the original protocol design can influence study results and hinder reproducibility.
Preregistration	Empirical studies† (n=132)	Preregistration can improve the transparency of paywall-restricted articles and unpublished trials.

		<p>Accessibility to methods, hypotheses, and analysis scripts can help mitigate reporting bias, data dredging, and p-hacking</p>
<p>Disclosure of Conflicts of Interest</p>	<p>All eligible studies‡ (n=290)</p>	<p>Authors disclosing all conflicts of interest that may compromise professional decision making (typically due to a secondary interest, such as financial gain) allows for publication of the most unbiased, authentic research possible.</p>
<p>Funding Source</p>	<p>All eligible studies‡ (n=290)</p>	<p>Sources of funding have an influential presence within the world of research. Funders help make costly study designs possible by contributing tangible resources to conduct experiments. Disclosure of funding contributions helps to increase the transparency of</p>

		medical research.
Open Access	All studies included in random sample§ (n=300)	Open access increases the availability of materials and protocols necessary for reproduction. Restricting access to such information precludes a comprehensive evaluation of a given study.
<p>* “Empirical studies” refers to studies with empirical data and includes the following classifications: clinical trial, cohort, case-control, secondary analysis, chart review, commentary (with data analysis), and cross-sectional; this category excludes case studies and case series due to the inherent difficulty surrounding reproducing these study types, as discussed by Wallach et al.<sup>22</sup> Meta-analyses were also excluded from this analysis as materials are not typically included.</p> <p>† Empirical studies (clinical trial, cohort, case-control, secondary analysis, chart review, commentary [with data analysis], and cross-sectional) excluding case reports and case series</p> <p>‡ Empirical studies (clinical trial, cohort, case-control, secondary analysis, chart review, commentary [with data analysis], and cross-sectional) including case reports and case series.</p> <p>§ All empirical and non-empirical studies were included in this analysis.</p>		

**Supplemental Table 2. Cardiovascular System Journal Impact Factor**

<b>Five-Year Impact Factor</b>	Median	4.023
	1st quartile	1.826
	3rd quartile	5.42
	Interquartile range	1st-3rd
<b>Most Recent Impact Factor Year (n=290)</b>		
<b>Most Recent Impact Factor Year (n=290)</b>	2014	0
	2015	0
	2016	0
	2017	264
	2018	20
	Not Found	16
<b>Most Recent Impact Factor</b>		
<b>Most Recent Impact Factor</b>	Median	3.78
	1st quartile	1.695
	3rd quartile	5.8
	Interquartile range	1st-3rd

**Supplemental Table 3. Additional Reproducibility and Transparency Characteristics of Publications in Cardiology Journals**

<b>Characteristic</b>	<b>Google Form Response</b>	<b>Response Rate, number (%)</b>	<b>95% CI</b>
<b>Data Availability Statement (n= 132)</b>	Statement present, the data (or some of the data) are available	32 (24)	0.183-0.277
	Statement declares the data are NOT available	1 (0.7)	-0.002-0.017
	No data availability statement present	99 (75.3)	0.714-0.811
<b>Means By Which Additional Data Are Available</b>	Personal or institutional website	0	-
	Supplementary information hosted by the journal	28	-
	Online third party repository	1	-
	Upon request from the corresponding author(s)	3	-
	Other	0	-
<b>Accessibility of Additional Data</b>	Data could be successfully accessed and downloaded	19	-
	Data could not be accessed and	13	-

	downloaded		
	Data files containing all raw data	6	-
	Data files without all raw data	13	-
<b>Materials Availability Statement (n= 132)</b>			
	Statement present, some materials are available	15 (11.4)	0.078-0.150
	Statement present, materials are NOT available	1 (0.8)	-0.002-0.174
	No materials availability statement present	116 (87.8)	0.842-0.916
<b>Means By Which Supplemental Materials Are Available</b>			
	Personal or institutional website	2	-
	Supplementary information hosted by the journal	10	-
	Online third party	2	-
	Upon request from the corresponding author(s)	1	-

<b>Accessibility of Additional Materials</b>	Materials availability statement present and materials were accessible	7	-
	Materials availability statement present, but the materials were not accessible	8	-
<b>Protocol Availability Statement (n= 132)</b>			
<b>Protocol Availability Statement (n= 132)</b>	Statement declaring full protocol was available	4 (3)	0.011-0.050
	No protocol availability statement present	128 (97)	0.950-0.989
<b>Accessibility of Additional Protocols</b>			
<b>Accessibility of Additional Protocols</b>	Protocol was NOT available using protocol link	2	-
	Protocol was available using to protocol link	2	-
	Hypotheses were included in the linked protocol	0	-
	Methods were included in the linked protocol	2	-
	Analysis plans were included in the linked protocol	2	-

<b>Analysis Script Availability Statement (n= 132)</b>	Statement present, declares that the analysis scripts (or some of the analysis scripts) are available	0 (0.0)	0.000-0.000
	Statement present, declares that the analysis scripts are NOT available	0 (0.0)	0.000-0.000
	No analysis script statement present	132 (100.0)	1
<b>Preregistration Statement (n=132)</b>	Statement declaring study was preregistered	6 (4.5)	0.022-0.069
	Statement declaring the study was NOT preregistered	1 (0.8)	-0.002-0.017
	No preregistration statement present	125 (94.7)	0.922-0.972
<b>Accessibility of Publication Registration</b>	Preregistration was accessible	4	-
	Preregistration was not accessible	2	-
	Number of studies	5	-

	preregistered on <a href="https://clinicaltrials.gov">ClinicalTrials.gov</a>		
	Number of studies pre-registered on UMIN Clinical Trials Registry (UMIN-CTR)	1	-
<b>Conflict of Interest Statement (n= 290)</b>			
	Disclosure statement declaring one or more conflicts of interest	64 (22)	0.174-0.268
	Disclosure statement declaring there are no conflicts of interest	177 (61)	0.555-0.666
	No conflict of interest statement present	49 (17)	0.127-0.211
<b>Open Access (n=300)</b>			
	Publication found via Open Access Button ( <a href="https://openaccessbutton.org">https://openaccessbutton.org</a> )	138 (46)	0.404-0.516
	Publication found via Google Scholar and/or PubMed	30 (10)	0.038-0.095
	Publication determined to be paywall restricted	132 (44)	0.384-0.496
Abbreviations- CI: Confidence Interval			



**Supplemental Table 4. Number of Times Publications Has Been Cited in a Meta-Analysis  
and/or Systematic Review Article**

	<b>n (%)</b>
<b>No citation</b>	106 (80.3)
<b>A Single Citation</b>	20 (15.2)
<b>One to Five Citations</b>	5 (3.8)
<b>Greater Than Five Citations</b>	1 (0.7)