bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright copyright under 17 USC 105 and is also made available for use under a CC0 license.

Research Report

Ethical Shades of Gray: Questionable Research Practices in Health Professions Education

Anthony R. Artino Jr, PhD, Erik W. Driessen, PhD, and Lauren A. Maggio, PhD

A.R. Artino Jr is professor of medicine, Department of Medicine, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland; ORCID: http://orcid.org/0000-0003-2661-7853; Twitter: @mededdoc

E.W. Driessen is professor of medical education, Faculty of Health, Medicine and Life Science, Maastricht University, the Netherlands. Twitter: @erikwdriessenclinical

L.A Maggio is associate professor of medicine, Department of Medicine, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland. ORCID: https://orcid.org/0000-0002-2997-6133; Twitter: @laurenmaggio

Correspondence should be addressed to Anthony R. Artino Jr, Department of Medicine, F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Rd., Bethesda, MD 20814; telephone: (301) 295-3693; e-mail: anthony.artino@usuhs.edu;

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, the U.S. Navy, the Department of Defense, or the U.S. Government.

Written work prepared by employees of the Federal Government as part of their official duties is, under the U.S. Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the United States Code is not available. As such, copyright does not extend to the contributions of employees of the Federal Government.

Word count = 3,608

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Abstract

Purpose

To maintain scientific integrity and engender public confidence, research must be conducted responsibly. Whereas scientific misconduct, like data fabrication, is clearly irresponsible and unethical, other behaviors—often referred to as questionable research practices (QRPs)—exploit the ethical shades of gray that color acceptable practice. This study aimed to measure the frequency of self-reported QRPs in a diverse, international sample of health professions education (HPE) researchers.

Method

In 2017, the authors conducted an anonymous, cross-sectional survey study. The web-based survey contained 43 QRP items that asked respondents to rate how often they had engaged in various forms of scientific misconduct. The items were adapted from two previously published surveys.

Results

In total, 590 HPE researchers took the survey. The mean age was 46 years (SD=11.6), and the majority of participants were from the United States (26.4%), Europe (23.2%), and Canada (15.3%). The three most frequently reported QRPs were adding authors to a paper who did not qualify for authorship (60.6%), citing articles that were not read (49.5%), and selectively citing papers to please editors or reviewers (49.4%). Additionally, respondents reported misrepresenting a participant's words (6.7%), plagiarizing (5.5%), inappropriately modifying

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

results (5.3%), deleting data without disclosure (3.4%), and fabricating data (2.4%). Overall, 533 (90.3%) respondents reported at least one QRP.

Conclusions

Notwithstanding the methodological limitations of survey research, these findings indicate that a substantial proportion of HPE researchers report a range of QRPs. In light of these results, reforms are needed to improve the credibility and integrity of the HPE research enterprise.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

> "Researchers should practice research responsibly. Unfortunately, some do not."
>
> –Nicholas H. Steneck, 2006[1]

The responsible conduct of research is the foundation of sound scientific practice.[1,2] The need to conduct research in a responsible manner is self-evident—if science is to inform our understanding of how the world works, it must be done in an honest, accurate, and unbiased way.[3]

Whereas behaviors like data fabrication are clearly irresponsible and highly unethical, other forms of research misconduct exploit the ethical shades of gray that color acceptable research practice. Often referred to as questionable research practices (QRPs), these behaviors "by nature of the very fact that they are often questionable as opposed to blatantly improper, also offer considerable latitude for rationalization and self-deception."[4] Consequently, QRPs are more prevalent and, many have argued, more damaging to science and its public reputation than obvious fraud.[4–8] Ultimately, QRPs can waste resources, provide an unfair advantage to some researchers over others, damage the scientific record, and provide a poor example for other researchers, especially trainees.[7]

Health professions education (HPE) is not immune to the damaging effects of irresponsible research practices. In the HPE context, we define QRPs as poor data management; inappropriate research procedures, including questionable procedures for obtaining informed consent; insufficient respect and care for study participants; improper research design; carelessness in observation and analysis; suboptimal trainee and mentor partnerships; unsuitable authorship or

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/252982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Gov Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

publishing practices; and derelictions in reviewing and editing.[7] The need to guard against such practices is frequently described in the author instructions for most scientific journals. For example, *Academic Medicine*'s author instructions clearly describe ethical considerations related to authorship, prior and duplicate publications, conflicts of interests, and ethical treatment of human subjects (http://journals.lww.com/academicmedicine/_layouts/15/1033/oaks.journals/informationforauthors.aspx). Such journal guidelines are often patterned on the recommendations of the Committee on Publication Ethics (https://publicationethics.org/).

In the last decade, commentaries by several HPE journal editors have highlighted instances of QRPs in article submissions, including self-plagiarism, so-called "salami slicing" (i.e., inappropriately dividing a single study into multiple papers), and unethical authorship practices.[9–11] Additionally, a recent review of four HPE journals found that 13% of original research articles published in 2013 did not address approval by an ethics review board or stated that it was unnecessary, without further discussion.[12] Moreover, a 2017 study of senior HPE leaders highlighted multiple problematic authorship practices, including honorary authorship and the exclusion of authors who deserved authorship.[13] Notably, only about half of the senior researchers surveyed were able to correctly identify the authorship standards used by most medical journals (i.e., the International Committee of Medical Journal Editors (ICMJE) authorship criteria[14]).

Notwithstanding these examples, QRPs have received limited attention in the HPE literature. In a recent article,[7] we attempted to raise the community's awareness of QRPs and highlight the need to examine their pervasiveness among HPE researchers. With this call to action in mind, we

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

conducted the present study to measure the frequency of self-reported QRPs in a diverse, international sample of HPE researchers. In doing so, we hope to continue the conversation about QRPs in our growing HPE field, with the ultimate goal of promoting the responsible conduct of high-quality, ethical research.

Method

To measure the frequency of serious research misconduct and other QRPs, many different approaches have been employed in the literature. These include counts of confirmed cases of researcher fraud and paper retractions, as well research audits by government funders.[8] Such methods are limited because they are calculated based on misconduct that has been discovered, and detecting such misconduct is difficult.[15] Moreover, distinguishing intentional misconduct from honest mistakes is challenging. Therefore, such approaches significantly underestimate the real frequency of QRPs, since only researchers know if they have willfully acted in a questionable or unethical manner.

To address these challenges, survey methods have been used to ask scientists directly about their research behaviors.[4–6,16,17] Like the measurement of any socially undesirable behavior, assessing QRPs via self-report likely underestimates the true prevalence or frequency of the behaviors. Nonetheless, when employed appropriately, survey methods can generate reasonable estimates that provide a general sense of the problem's scope.[18,19]

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Therefore, we administered an anonymous, cross-sectional survey to determine the frequency of QRPs in a sample of HPE researchers. The Ethical Review Board Committee of the Netherlands Association for Medical Education approved this study (Dossier #937).

Survey development

We developed our survey instrument by adapting several existing surveys. The final version of the survey featured a total of 66 items divided into three sections (see Supplemental Digital Appendix). The first section included 43 items derived from two previously published surveys assessing QRPs in biomedicine.[5,6] The QRP items asked respondents to identify how often they had engaged in the particular research practice. The practices spanned the research continuum, from data collection and storage to study reporting, collaboration, and authorship. The QRP items employed a six-point, Likert-type, frequency-response scale: *never*, *once*, *occasionally*, *sometimes*, *frequently*, and *almost always*. Each item also included the response option *not applicable to my work*.

We slightly modified the original QRP items to improve their clarity and relevance to the HPE research context. For example, the original item "inadequately handled or stored data or (bio)materials" was revised to "inappropriately stored sensitive research data (e.g., data that contains personally identifiable information)." Following these modifications, 19 experienced HPE researchers reviewed the adapted survey items and provided detailed qualitative feedback.[20] Ten of the expert reviewers were women, and all held doctoral degrees (13 PhDs, 4 MD/PhDs, and 2 MDs). On average, the reviewers had published 98.9 journal articles (SD=66.1) findable in PubMed. Based on the date of their first HPE publication, they had been publishing in the field

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

for an average of 20.7 years (SD=9.3). Expert reviewers reported their work location as the United States (n=9), Canada (n=3), Europe (n=2), South America (n=1), Africa (n=1), and Australia (n=1), and all but two were identified as full professors.

The expert feedback included comments on item relevance, clarity, missing facets, and suggestions for overall survey improvement. Based on the expert feedback, we revised the survey again; revisions included wording modifications and the development of several new items specific to HPE research. For example, based on the recommendations of several experts, we created the following items related to qualitative research methods, among several others: "misrepresented a participant's words or writings" and "claimed you used a particular qualitative research approach appropriately (e.g., grounded theory) when you knowingly did not."

The second section of the survey included nine publication pressure items[16] (these data are not reported here), and the final section included 13 demographic items. The survey ended with a single, open-ended question, which gave respondents the opportunity to give feedback on the survey instrument itself or provide additional thoughts on QRPs in HPE.

Sampling and survey distribution procedures

To create our sample, we used two separate approaches. First, we created a "curated sample" by searching Web of Science, Scielo (a database focused on South America), African Journals Online, and Asia Journals Online for articles in over 20 HPE journals published in 2015 and 2016. From these articles, we extracted all available author email addresses, removing duplicate authors. This process generated a sample of 1,840 unique HPE researchers. All names and emails

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/252982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

were entered into Qualtrics, an online survey tool (Qualtrics, Provo, Utah), and the survey was then distributed in four waves of email invitations: wave 1 (sent November 13, 2017), wave 2 (sent November 20, 2017), wave 3 (sent November 27, 2017), and wave 4 (sent December 11, 2017).

Next, we collected a "social media sample" by posting anonymous links to the survey on our Twitter and Facebook accounts (posted on December 11, 2017). All survey responses obtained from the social media links were tracked separately from those sent to the curated sample. To prevent duplicate submissions, respondents in the social media sample were given the option to select "I have already completed this survey" on the informed consent page.

Statistical analyses

Prior to analysis, we screened the data for accuracy and missing values. Next, we calculated the response rate in the curated sample using response rate definition #6, as delineated by the American Association for Public Opinion Research.[21] Then, to assess potential nonresponse bias in the curated sample, we used wave analysis to calculate a nonresponse bias statistic.[22] In wave analysis, late respondents are considered proxies for non-responders, and their responses are compared to responses from the first wave. In addition, to determine whether or not it was appropriate to combine the curated and social media samples for analysis, we conducted a multivariate analysis of variance (MANOVA) to compare respondents on several demographic characteristics: age, experience doing HPE research, percentage of work time dedicated to HPE research, and number of peer-reviewed publications. Finally, we calculated descriptive statistics for the total sample, with particular emphasis on the frequency of self-reported QRPs. All data

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

analyses and visualizations were conducted using IBM SPSS Statistics (IBM Corporation, New York, NY) and Microsoft Excel (Microsoft Corporation, Redmond, WA), respectively.

Results

Of the 1,840 email invitations sent to HPE researchers in the curated sample, 199 were returned as undeliverable, leaving 1,641 potential respondents. Of these, 463 (28.2%) researchers completed a least a portion of the survey.[21] Results from the wave analysis revealed a nonresponse bias statistic of 0.36. On a six-point, frequency-response scale, this represents a 6% difference, which is unlikely to have a meaningful effect on practical interpretation of the results.[22]

The social media sample yielded an additional 127 responses. Results from the MANOVA comparing respondents in the curated sample to those in the social media sample revealed statistically significant differences between the two groups $F(5, 524) = 6.67$, $P < .001$. In particular, post-hoc analyses revealed that respondents in the social media sample were slightly younger (M=40.7 years) and more inexperienced in HPE research (M=7.5 years) than those in the curated sample (M=47.4 years and M=11.0 years, respectively). That said, the two groups did not differ in terms of percentage of work time spent doing HPE research activities or mean number of peer-reviewed publications. Therefore, because our goal was to explore the frequency of QRPs among a diverse, international sample of HPE researchers, we elected to pool the two samples and analyze the data together.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Of the 590 respondents in the pooled sample, the mean age was 46 years (SD=11.6), and there were 305 (51.7%) women, 246 (41.7%) men, and 39 (6.6%) individuals who did not report their gender. As indicated in Table 1, the sample consisted of HPE researchers from across the World Health Organization's six world regions. The majority reported their location as the United States (26.4%), Europe (23.2%), and Canada (15.3%). Respondents' education, area of study, work context and role, academic rank, and primary research activities are also presented in Table 1. In addition, respondents reported the following: years involved in HPE in any capacity (M=14.9 years, SD=9.7), years involved in HPE research (M=11.3 years, SD=8.5), percentage of work time spent conducting HPE research (M=27.3%, SD=23.7%), and total number of peer-reviewed publications (M=40.1, SD=55.0).

Table 2 summarizes the frequency of self-reported QRPs, and the Figure provides a visual representation of these results. To simplify the figure, we collapsed the response options of *occasionally*, *sometimes*, *frequently*, and *almost always* into a single frequency option labeled *more than once*. Finally, the Box lists the top 10 most frequently reported QRPs among our respondents, from highest to lowest frequency. As indicated, the most frequently reported QRPs were related to authorship and study reporting practices, as well as issues around data storage, collection, and interpretation. Additionally, 39 (6.7%) respondents reported misrepresenting a participant's words, 31 (5.5%) reported using sections of text from another author's copyrighted material without permission or proper citation, 30 (5.3%)reported inappropriately modifying study results due to pressure from a research advisor or collaborator, 20 (3.4%) reported deleting data before performing analysis without disclosure, and 14 (2.4%) reported fabricating data. Overall, 533 (90.3%) respondents reported at least one QRP.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Box The top 10 most frequently reported QRPs among an international sample of 590 health professions education researchers (listed from highest to lowest frequency).

1. Added one or more authors to a paper who did not qualify for authorship (so-called "honorary authorship")
2. Cited articles and or materials that you have not read
3. Selectively cited certain papers just to please editors or reviewers
4. Inappropriately stored sensitive research data (e.g., data that contains personally identifiable information)
5. Selectively cited your own work just to improve your citation metrics
6. Ignored a colleague's questionable interpretation of data
7. Collected course or curriculum data under the guise of "program evaluation" without human-subjects ethics (IRB) approval with the ultimate intent of using the data for research purposes
8. Inappropriately emailed sensitive research data (e.g., data that contains personally identifiable information)
9. Accepted authorship for which you did not qualify (so-called "honorary authorship")
10. Spread study results over more papers than is appropriate (so-called "salami slicing")

Discussion

This study examined the frequency of self-reported QRPs among HPE researchers, practices that may be detrimental to scientific inquiry.[1,2] To our knowledge, this is the first study to explore QRPs across the HPE research continuum. Taken together, our findings indicate that a substantial proportion of HPE researchers admit to having engaged in a range of QRPs. These results are consistent with the extant literature on research misconduct in other fields,[4–6,8,16] and they are important because QRPs can waste resources, provide an unfair advantage to some researchers over others, and ultimately impede scientific progress. Therefore, this study raises significant concerns about the credibility of HPE research, suggesting that our community may need to take a hard look at its ethical norms and research culture.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

In our survey, we asked respondents about a range of problematic behaviors, from clear misconduct (data fabrication) and falsification (distortion of results) to plagiarism (copying ideas or words without attribution) and authorship manipulation (e.g., honorary authorship). As Fanelli[8] noted in his review of research misconduct, certain activities (e.g., honorary authorship or excluding study limitations) are qualitatively different than fabrication and falsification because they do not *directly* distort the quality of the science, per se. However, the damage done to the scientific enterprise by these "less severe" and more ambiguous QRPs may be proportionally greater than deliberate misconduct, for the simple reason that such practices occur more frequently. For example, 20.1% of HPE respondents reported one or more instances of "salami slicing." While some may think of this as a minor offense, the practice fills the literature with more articles than is seemingly necessary.[1,10] So, not only does this activity unfairly reward authors and waste resources (e.g., editorial time and journal space), it also can inflate the significance of a given finding, which in turn can distort the outcomes of meta-analyses and other types of systematic reviews.[10,23]

It is worth stating that the interpretation of our results is limited by the nature of the survey methodology employed and, in particular, by the threat of nonresponse bias (especially considering the sensitive nature of the topic under study).[5,6] Therefore, it is reasonable to ask: (1) how reliable are these frequency estimates, and (2) what can they really say about the actual frequency of QRPs among HPE researchers? Although we did not assess score reliability in the present study, we argue here, as others have previously,[5,8,16] that self-reports of QRPs likely underestimate the real frequency of questionable behaviors. Researchers who have acted unethically are undoubtedly hesitant to reveal such activities in a survey, despite all assurances

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

of anonymity. What is more, the opposite—researchers admitting to unethical or otherwise questionable practices that they did not do—seems unlikely.[8] Therefore, we speculate that QRPs may be even more widespread in our community than our estimates imply. Nevertheless, rather than establishing an absolute prevalence of QRPs in HPE, we believe these data are better suited for helping the community understand the nature of the most common QRPs and begin finding feasible solutions to improve our research enterprise.

Questionable research practices related to authorship were some of the most frequently reported behaviors in this study, particularly the practice of giving or accepting unwarranted authorship (so-called "honorary authorship"). Honorary authorship is unethical in academic publishing because individuals who have not sufficiently contributed to the work unfairly receive credit as an author and misrepresent their contributions in the scientific literature.[24] Our findings corroborate the results of a recent survey of established HPE researchers,[13] and it seems we are not alone in these practices.[24–26] For example, a 2008 study of six high-impact medical journals found that 17.6% of corresponding authors admitted to including honorary authors.[24] In a separate survey of radiology researchers, 58.9% of respondents reported that they had written a paper with a co-author whose contributions did not merit authorship.[27] In some fields, including HPE, this practice has led journals to require authors to sign an author contribution agreement to verify their explicit authorship roles.[28] The effectiveness of such requirements is unknown and could be a fruitful area for additional research.

Of note in our findings, the frequency of authors giving honorary authorship and those accepting honorary authorship were not equivalent: 60.6% admitted to adding undeserving authors whereas

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

only 22.7% admitted to accepting honorary authorship. This mismatch suggests that HPE researchers may not fully understand authorship criteria. It may also be a concrete example of the so-called "Muhammad Ali effect"—the idea that individuals often see themselves as more likely to perform good acts and less likely to perform bad acts than others.[29] Regardless of the mechanism, this finding indicates the need for increased author communication and better shared understanding of author roles and responsibilities, such as those set forth by the ICMJE (http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html).

A complete discussion of all the QRPs assessed in this study is outside the scope of this paper. While readers may debate the degree to which some of these behaviors are unethical or otherwise problematic, we should note, as described above, that seemingly minor infractions can have far-reaching negative consequences. For example, by employing QRPs like p-hacking (i.e., manipulating data or analyses until nonsignificant results become significant)[30,31] or taking advantage of other types of "researcher degrees of freedom,"[32] scientists can discover "illusory results"[33] that actually represent artifacts of their study design and analytic approach, as opposed to legitimate findings that can be replicated.[34,35] Some have argued that such behaviors are the result of individual researchers responding to a set of incentives, the most important of which are rewards for the *quantity* (not the quality) of their publications.[16,36] A compelling way to view these behaviors is through the evolutionary lens of natural selection. Smaldino and McElreath[37] made this point, contending that "The persistence of poor methods results partly from incentives that favor them, leading to the natural selection of bad science. This dynamic requires no conscious strategizing—no deliberate cheating nor loafing—by scientists, only that publication is

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

a principal factor for career advancement." If one accepts this thesis, then the most effective way to improve research practices, and the quality of the corresponding science, is to change incentive structures at the institutional level.[37]

Limitations and future directions

The current study has limitations. First, nonresponse bias is a legitimate concern, especially considering the modest response rate (28.2%) in the curated sample. That said, recent research suggests that response rate may be a flawed indicator of response quality and representativeness.[38,39] Moreover, the wave analysis results indicate that nonresponse bias was limited in our sample. Nonetheless, investigators should build on these initial results by examining QRPs among a larger, more global sample of HPE researchers.

A second limitation relates to the inherent challenge of assessing complex, context-specific research behaviors with a survey that requires respondents to self-assess their own practices.[40,41] Because some of the QRPs on our survey are judgment calls, their evaluation likely requires more detail and nuance than an individual survey item can provide.[40] So, while the practice of fabricating data is fairly straightforward (and never justified), the same cannot be said for something like inappropriately storing sensitive research data. The latter practice is open to interpretation: what is considered "inappropriate storage" to one researcher might seem perfectly fine to another. We attempted to reduce this type of ambiguity in our survey items by employing a rigorous expert review process, which resulted in several revised (and we believe, improved) survey items. But, ultimately, "survey self-reports can never fully rule out ambiguities in meaning, limitations in autobiographical memory, or motivated biases."[40] Thus, future work

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

might apply qualitative research methods to address some of these limitations and further unpack the nature of QRPs in HPE.

Finally, we administered our survey in English and did not ask respondents to focus on a particular time period. These implementation and design choices could have negatively affected data quality. For example, several respondents noted in their written comments that ethical standards related to human-subjects research had evolved over time. Future research could address this problem by limiting the time period that respondents are asked to consider.[5]

Recommendations for practice

As HPE continues to mature as a field, it is essential that we explicitly confront our obligation to conduct our research in an ethical and responsible manner. Previously, we suggested a number of recommendations to improve practice,[7] and the findings reported here indicate the time is *now* to implement these and other policy changes. Our recommended approaches included: (1) empowering research mentors as role models, (2) openly airing research dilemmas and infractions, (3) protecting whistleblowers, (4) establishing mechanisms for facilitating responsible research (e.g., creating HPE-specific institutional review boards[42]), and (5) rewarding responsible researchers (e.g., providing grant funding and publication opportunities for replication studies[34]). These are institution-level approaches; they embrace the idea that QRPs are not simply the result of individual researchers acting badly. Instead, there are important contextual factors that influence researcher behavior (e.g., social norms, power disparities, institutional policies, and academic incentives).[7] Examples of initiatives being tried in other fields and institutions include promotion and tenure guidelines that privilege publication

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

quality over quantity,[43] study pre-registration plans,[33] and other open-science practices (e.g., open data and materials sharing).[44] All of these approaches require study to determine their efficacy in HPE.

In summary

Cultivating the responsible conduct of research is essential if we are to maintain scientific integrity and engender public confidence in our research.[1,2] This study raises significant concerns about the credibility of HPE research and presents a somewhat pessimistic picture of our community. We should be clear though—most HPE scientists surveyed did *not* report the vast majority of QRPs. They are presumably doing good science. Nevertheless, we believe reforms are needed. In addition, we recommend future research to monitor QRPs in HPE and evaluate the effectiveness of policies designed to improve the integrity of our research enterprise.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/252982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright copyright under 17 USC 105 and is also made available for use under a CC0 license.

Acknowledgments: The authors wish to sincerely thank all of the HPE researchers who took the time (and had the courage) to complete our survey.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: The Ethical Review Board Committee of the Netherlands Association for Medical Education approved this study (Dossier #937).

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, the U.S. Navy, the Department of Defense, or the U.S. Government.

Written work prepared by employees of the Federal Government as part of their official duties is, under the U.S. Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the United States Code is not available. As such, copyright does not extend to the contributions of employees of the Federal Government.

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

References

1. Steneck NH. Fostering integrity in research: definitions, current knowledge, and future directions. *Sci Eng Ethics*. 2006;12(1):53-74.
2. Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature*. 2005;435(7043):737-738. doi:10.1038/435737a.
3. Steneck NH. Introduction to the Responsible Conduct of Research. https://ori.hhs.gov/sites/default/files/rcrintro.pdf. Accessed January 15, 2018.
4. John LK, Loewenstein G, Prelec D. Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychol Sci*. 2012;23(5):524-532. doi:10.1177/0956797611430953.
5. Tijdink JK, Bouter LM, Veldkamp CLS, van de Ven PM, Wicherts JM, Smulders YM. Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. Dorta-González P, ed. *PLoS One*. 2016;11(9):e0163251. doi:10.1371/journal.pone.0163251.
6. Bouter LM, Tijdink J, Axelsen N, Martinson BC, ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Res Integr Peer Rev*. 2016;1(1):17. doi:10.1186/s41073-016-0024-5.
7. Maggio LA, Artino AR, Picho K, Driessen EW. Are You Sure You Want to Do That? Fostering the Responsible Conduct of Medical Education Research. *Acad Med*. July 2017:1. doi:10.1097/ACM.0000000000001805.
8. Fanelli D. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. Tregenza T, ed. *PLoS One*. 2009;4(5):e5738.

bioRxiv preprint first posted online Jan 31, 2018; doi:http://dx.doi.org/10.1101/252982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

doi:10.1371/journal.pone.0005738.

9. Brice J, Bligh J, Bordage G, et al. Publishing ethics in medical education journals. *Acad Med*. 2009;84(10 Suppl):S132-4. doi:10.1097/ACM.0b013e3181b36f69.

10. Eva KW. How would you like your salami? A guide to slicing. *Med Educ*. 2017;51(5):456-457. doi:10.1111/medu.13285.

11. ten Cate O. Why the ethics of medical education research differs from that of medical research. *Med Educ*. 2009;43(7):608-610. doi:10.1111/j.1365-2923.2009.03385.x.

12. Hally E, Walsh K. Research ethics and medical education. *Med Teach*. 2016;38(1):105-106. doi:10.3109/0142159X.2014.956068.

13. Uijtdehaage S, Mavis B, Durning SJ. Whose paper is it anyway? Authorship criteria according to established scholars in health professions education. *Acad Med*.

14. International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. 2017. http://www.icmje.org/icmje-recommendations.pdf. Accessed January 15, 2018.

15. White C, ed. *The COPE Report 2000: Annual Report of the Committee on Publication Ethics*. London, UK: BMJ Books; 2000.

16. Tijdink JK, Verbeke R, Smulders YM. Publication Pressure and Scientific Misconduct in Medical Scientists. *J Empir Res Hum Res Ethics*. 2014;9(5):64-71. doi:10.1177/1556264614552421.

17. Anderson MS, Horn AS, Risbey KR, Ronning EA, De Vries R, Martinson BC. What Do Mentoring and Training in the Responsible Conduct of Research Have To Do with Scientists' Misbehavior? Findings from a National Survey of NIH-Funded Scientists. *Acad Med*. 2007;82(9):853-860. doi:10.1097/ACM.0b013e31812f764c.

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

18. Schaeffer NC, Dykema J. Questions for Surveys: Current Trends and Future Directions. *Public Opin Q*. 2011;75(5):909-961. doi:10.1093/poq/nfr048.

19. Tourangeau R, Yan T. Sensitive questions in surveys. *Psychol Bull*. 2007;133(5):859-883. doi:10.1037/0033-2909.133.5.859.

20. Artino Jr. AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach*. 2014;36:463-474. doi:10.3109/0142159x.2014.889814.

21. The American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th ed. AAPOR; 2016.

22. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE Guide No. 102. *Med Teach*. 2016;38(3):217-228. doi:10.3109/0142159X.2015.1105945.

23. Supak Smolcić V. Salami publication: definitions and examples. *Biochem medica*. 2013;23(3):237-241. doi:10.11613/BM.2013.030.

24. Wislar JS, Flanagin A, Fontanarosa PB, Deangelis CD. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. *BMJ*. 2011;343:d6128.

25. Kornhaber RA, McLean LM, Baber RJ. Ongoing ethical issues concerning authorship in biomedical journals: an integrative review. *Int J Nanomedicine*. 2015;10:4837-4846. doi:10.2147/IJN.S87585.

26. Vera-Badillo FE, Napoleone M, Krzyzanowska MK, et al. Honorary and ghost authorship in reports of randomised clinical trials in oncology. *Eur J Cancer*. 2016;66:1-8. doi:10.1016/J.EJCA.2016.06.023.

27. Eisenberg RL, Ngo L, Boiselle PM, Bankier AA. Honorary Authorship in Radiologic

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Research Articles: Assessment of Frequency and Associated Factors. *Radiology*. 2011;259(2):479-486. doi:10.1148/radiol.11101500.

28. Lundberg GD, Flanagin A. New Requirements for Authors: Signed Statements of Authorship Responsibility and Financial Disclosure. *JAMA J Am Med Assoc*. 1989;262(14):2003. doi:10.1001/jama.1989.03430140121037.

29. Allison ST, Messick DM, Goethals GR. On Being Better but not Smarter than Others: The Muhammad Ali Effect. *Soc Cogn*. 1989;7(3):275-295. doi:10.1521/soco.1989.7.3.275.

30. 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. doi:10.1371/journal.pbio.1002106.

31. Nuzzo R. Scientific method: Statistical errors. *Nature*. 2014;506(7487):150-152. doi:10.1038/506150a.

32. Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci*. 2011;22(11):1359-1366. doi:10.1177/0956797611417632.

33. Gehlbach H, Robinson CD. Mitigating Illusory Results through Preregistration in Education. *J Res Educ Eff*. October 2017:1-20. doi:10.1080/19345747.2017.1387950.

34. Picho K, Maggio LA, Artino AR. Science: the slow march of accumulating evidence. *Perspect Med Educ*. 2016;5(6):350-353. doi:10.1007/s40037-016-0305-1.

35. Ioannidis JPA. Why Most Published Research Findings Are False. *PLoS Med*. 2005;2(8):e124. doi:10.1371/journal.pmed.0020124.

36. Horton R. Offline: What is medicine's 5 sigma? *Lancet*. 2015;385(9976):1380. doi:10.1016/S0140-6736(15)60696-1.

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

37. Smaldino PE, McElreath R. The natural selection of bad science. *R Soc Open Sci*. 2016;3(9):160384. doi:10.1098/rsos.160384.

38. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. *JAMA*. 2012;307:1805-1806. doi:10.1001/jama.2012.3532.

39. Halbesleben JRB, Whitman M V. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. *Health Serv Res*. 2013;48(3):913-930. doi:10.1111/1475-6773.12002.

40. Fiedler K, Schwarz N. Questionable Research Practices Revisited. *Soc Psychol Personal Sci*. 2016;7(1):45-52. doi:10.1177/1948550615612150.

41. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med*. 2005;80(10 Suppl):S46-54. http://www.ncbi.nlm.nih.gov/pubmed/16199457.

42. DeMeo SD, Nagler A, Heflin MT. Development of a Health Professions Education Research-Specific Institutional Review Board Template. *Acad Med*. 2016;91(2):229-232. doi:10.1097/ACM.0000000000000987.

43. Nazim Ali S, Young HC, Ali NM. Determining the quality of publications and research for tenure or promotion decisions. *Libr Rev*. 1996;45(1):39-53. doi:10.1108/00242539610107749.

44. Kidwell MC, Lazarević LB, Baranski E, et al. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. Macleod MR, ed. *PLOS Biol*. 2016;14(5):e1002456. doi:10.1371/journal.pbio.1002456.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Table 1

Demographic characteristics of an international sample of 590 health professions education researchers

Demographic characteristic		No. (%)
Gender	Female	305 (51.7)
	Male	246 (41.7)
	Not reported	39 (6.6)
Region	United States	156 (26.4)
	Europe (not including United Kingdom)	96 (16.3)
	Canada	90 (15.3)
	Australia and New Zealand	42 (7.1)
	United Kingdom	41 (6.9)
	Africa	37 (6.3)
	Asia	37 (6.3)
	South/Latin America and Caribbean	18 (3.0)
	Middle East	16 (2.7)
	Other	12 (2.0)
	Not reported	45 (7.6)
Education [a]	Bachelor's degree	245 (44.4)
	Master's degree	306 (55.5)
	MD/DO degree	276 (50.0)
	PhD/EdD degree	323 (58.5)
	Other professional degree	30 (5.4)
Area of study for highest degree selected	Social science	293 (49.7)
	Clinical science	167 (28.3)
	Other	42 (7.1)
	Basic science	39 (6.6)
	Humanities	11 (1.9)
	Not reported	38 (6.4)
Work context [a]	Undergraduate medical education	353 (65.7)
	Graduate medical education	308 (57.4)
	Continuing medical education	142 (26.4)
Work role	Researcher	174 (29.5)
	Clinician	136 (23.1)
	Administrator/program director	89 (15.1)
	Teacher	86 (14.6)
	Other	65 (11.0)
	Not reported	40 (6.8)
Academic rank	Professor	140 (23.7)
	Associate professor	109 (18.5)
	Assistant professor	99 (16.8)
	Lecturer/instructor	64 (10.9)
	Graduate student	42 (7.1)
	Postdoctoral fellow	23 (3.9)
	Resident	16 (2.7)
	Other	52 (8.8)
	Not reported	45 (7.6)
Primary research activity	Mixed-method	280 (47.5)
	Quantitative	149 (25.3)
	Qualitative	119 (20.2)
	Not reported	42 (7.1)

Abbreviations: MD, Doctor of Medicine; DO, Doctor of Osteopathic Medicine.

[a] Percentages total more than 100% because respondents could select more than one category.

bioRxiv preprint first posted online April 12, 2018; doi: http://doi.org/10.1101/253988. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Table 2

Frequency of self-reported questionable research practices among an international sample of 590 health professions education researchers, reported as No. (%)[a]

In your work as an HPE researcher, how often have you engaged in any of the following behaviors, even if it has been only on a single occasion? If applicable, please consider your experiences with both quantitative and qualitative research.

	N	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable
Data Collection and Storage								
1-Conducted a human-subjects research study without ethics approval (i.e., without institutional review board [IRB] approval)	589	465 (78.9)	50 (8.5)	38 (6. 5)	13 (2.2)	6 (1.0)	4 (0.7)	13 (2.2)
2-Circumvented one or more aspects of human-subjects ethics rules (i.e., IRB rules)	585	463 (79.1)	47 (8.0)	49 (8.4)	10 (1.7)	4 (0.7)	0 (0.0)	12 (2.1)
3-Collected course or curriculum data under the guise of "program evaluation" without human-subjects ethics (IRB) approval with the ultimate intent of using the data for research purposes	588	420 (71.4)	45 (7.7)	65 (11.1)	24 (4.1)	9 (1.5)	2 (0.3)	23 (3.9)
4-Inappropriately stored sensitive research data (e.g., data that contains personally identifiable information)	585	353 (60.3)	56 (9.6)	119 (20.3)	36 (6.2)	15 (2.6)	2 (0.3)	4 (0.7)
5-Inappropriately emailed sensitive research data (e.g., data that contains personally identifiable information)	585	436 (74.5)	36 (6.2)	84 (14.4)	20 (3.4)	3 (0.5)	0 (0.0)	6 (1.0)
6-Stopped collecting data earlier than planned because the results already reached statistical significance, without formal stopping rules	586	497 (84.8)	22 (3.8)	13 (2.2)	6 (1.0)	3 (0.5)	1 (0.2)	44 (7.5)
7-Fabricated data	586	570 (97.3)	10 (1.7)	2 (0.3)	1 (0.2)	0 (0.0)	1 (0.2)	2 (0.3)
8-Pressured a student or other subordinate to be a study participant in your research	583	536 (91.9)	14 (2.4)	20 (3.4)	4 (0.7)	1 (0.2)	0 (0.0)	8 (1.4)
9-Used students or residents as research subjects without informing the overseeing dean, program director, or other pertinent official	584	468 (80.1)	23 (3.9)	46 (7.9)	17 (2.9)	8 (1.4)	2 (0.3)	20 (3.4)
Data Analysis								
	N	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable
10-Deleted data before performing data analysis without disclosure	582	553 (95.0)	9 (1.5)	11 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	9 (1.5)
11-Ignored a colleague's use of flawed data	581	447 (76.9)	36 (6.2)	46 (7.9)	16 (2.8)	5 (0.9)	1 (0.2)	30 (5.2)
12-Ignored a colleague's questionable interpretation of data	581	404 (69.5)	44 (7.6)	88 (15.1)	17 (2.9)	2 (0.3)	2 (0.3)	24 (4.1)
13-Reported a downwardly rounded p-value (e.g., reporting that a p-value of .054 is less than .05)	581	513 (88.3)	11 (1.9)	8 (1.4)	0 (0.0)	0 (0.0)	0 (0.0)	49 (8.4)
14-Misrepresented a participant's words or writings	580	526 (90.7)	16 (2.8)	21 (3.6)	2 (0.3)	0 (0.0)	0 (0.0)	15 (2.6)
15-Decided whether to exclude non-outlier data after looking at the impact of doing so on the results	584	470 (80.5)	32 (5.5)	32 (5.5)	7 (1.2)	1 (0.2)	0 (0.0)	42 (7.2)
16-In a qualitative study, failed to report disconfirming examples or cases that weaken your conclusions	583	422 (72.4)	32 (5.5)	32 (5.5)	6 (1.0)	1 (0.2)	0 (0.0)	90 (15.4)
17-Collected more data after seeing that the results were almost statistically significant	583	434 (74.4)	33 (5.7)	40 (6.9)	4 (0.7)	3 (0.5)	1 (0.2)	68 (11.7)
18-To confirm a hypothesis, selectively deleted or changed data after performing data analysis	583	523 (89.7)	16 (2.7)	6 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)	38 (6.5)
19-Reported an unexpected finding as having been hypothesized from the start	583	433 (74.3)	47 (8.1)	55 (9.4)	9 (1.5)	4 (0.7)	2 (0.3)	33 (5.7)
20-Concealed results that contradicted your previous findings or convictions	583	538 (92.3)	23 (3.9)	11 (1.9)	1 (0.2)	0 (0.0)	0 (0.0)	10 (1.7)
21-Claimed you used a particular qualitative research approach appropriately (e.g., grounded theory) when you knowingly did not	566	432 (76.3)	39 (6.9)	22 (3.9)	3 (0.5)	0 (0.0)	0 (0.0)	70 (12.4)

bioRxiv preprint first posted online April 12, 2018; doi: http://dx.doi.org/10.1101/253988. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 License.

22-Claimed you used a particular qualitative research technique appropriately (e.g., saturation, triangulation) when you knowingly did not	566	433 (76.5)	42 (7.4)	18 (3.2)	3 (0.5)	0 (0)	0 (0)	70 (12.4)
Study Reporting								
	N	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable
23-Spread study results over more papers than is appropriate (so-called "salami slicing")	566	442 (78.1)	59 (10.4)	52 (9.2)	2 (0.4)	1 (0.2)	0 (0.0)	10 (1.8)
24-Deliberately failed to mention important limitations of a study in the published paper	566	517 (91.3)	19 (3.4)	26 (4.6)	3 (0.5)	0 (0.0)	0 (0.0)	1 (0.2)
25-Deliberately failed to mention an organization that funded your research in the published paper	565	549 (97.2)	5 (0.9)	1 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	10 (1.8)
26-Inappropriately modified the results of a study due to pressure from a research advisor or other collaborator	564	531 (94.1)	21 (3.7)	8 (1.4)	1 (0.2)	0 (0.0)	0 (0.0)	3 (0.5)
27-Inappropriately modified the results of a study due to pressure from a funding agency	566	532 (94.0)	5 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	29 (5.1)
28-Failed to disclose relevant financial or intellectual conflicts of interest	567	548 (96.6)	6 (1.1)	4 (0.7)	1 (0.2)	0 (0.0)	0 (0.0)	8 (1.4)
29-Used someone else's ideas without their permission or proper citation	567	536 (94.5)	15 (2.6)	14 (2.5)	1 (0.2)	0 (0.0)	0 (0.0)	1 (0.2)
30-Used sections of text from another author's copyrighted material without permission or proper citation	566	534 (94.3)	13 (2.3)	16 (2.8)	2 (0.4)	0 (0.0)	0 (0.0)	1 (0.2)
31-Used sections of text from your own publications without proper citation (so-called "self-plagiarism")	564	449 (79.6)	35 (6.2)	65 (11.5)	12 (2.1)	0 (0.0)	0 (0.0)	3 (0.5)
32-Selectively cited certain papers just to please editors or reviewers	566	284 (50.2)	57 (10.1)	170 (30.0)	39 (6.9)	7 (1.2)	7 (1.2)	2 (0.4)
33-Cited articles and or materials that you have not read	565	284 (50.3)	39 (6.9)	181 (32.0)	48 (8.5)	12 (2.1)	0 (0.0)	1 (0.2)
34-Selectively cited your own work just to improve your citation metrics	562	391 (69.6)	28 (5.0)	101 (18.0)	24 (4.3)	15 (2.7)	2 (0.4)	1 (0.2)
35-Reused previously published data without disclosure (co-called "duplicate publication")	564	545 (96.6)	11 (2.0)	4 (0.7)	1 (0.2)	0 (0.0)	0 (0.0)	3 (0.5)
36-Used confidential information obtained as a reviewer or editor for your own research or publications	564	543 (96.3)	11 (2.0)	6 (1.1)	0 (0.0)	0 (0.0)	0 (0.0)	4 (0.7)
Collaboration and Authorship								
	N	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable
37-Refused to share data with legitimate colleagues	564	535 (94.9)	11 (2.0)	6 (1.1)	0 (0.0)	2 (0.4)	2 (0.4)	8 (1.4)
38-Added one or more authors to a paper who did not qualify for authorship (so-called "honorary authorship")	563	219 (38.9)	91 (16.2)	163 (29.0)	54 (9.6)	28 (5.0)	5 (0.9)	3 (0.5)
39-Accepted authorship for which you did not qualify (so-called "honorary authorship")	564	433 (76.8)	78 (13.8)	43 (7.6)	6 (1.1)	1 (0.2)	0 (0.0)	3 (0.5)
40-Demanded authorship for which you did not qualify (so-called "honorary authorship")	565	550 (97.3)	8 (1.4)	3 (0.5)	1 (0.2)	0 (0.0)	0 (0.0)	3 (0.5)
41-Omitted a contributor who deserved authorship	563	531 (94.3)	23 (4.1)	7 (1.2)	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.4)
42-Submitted (or re-submitted) a manuscript or grant application without consent from one or more of the authors	563	499 (88.6)	27 (4.8)	29 (5.2)	4 (0.7)	2 (0.4)	0 (0.0)	2 (0.4)
43-Submitted the same manuscript to multiple journals at once (so-called "duplicate" or "double submission")	564	558 (98.9)	4 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.4)

Questionable research practices are listed in the order presented on the survey.

[a] Percentages are calculated using data from individuals who responded to the item (i.e., non-responder data are not included in the denominator).

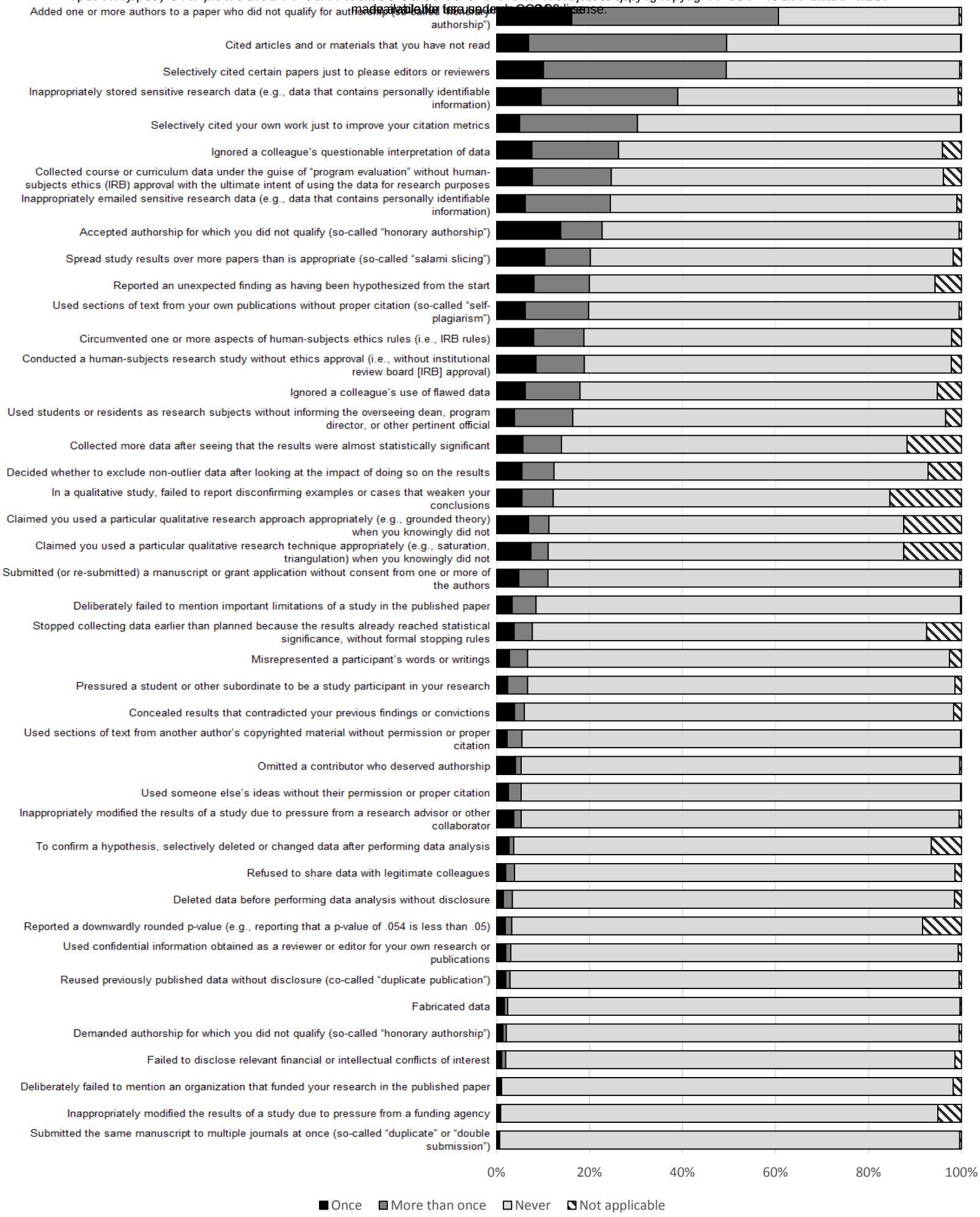
bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Figure Legend

Figure

Stacked bar graph showing the frequency of self-reported questionable research practices among an international sample of 590 health professions education researchers.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/252982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under CC BY-NC-ND license.



bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982; this version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Questionable Research Practices Survey

Thank you for your interest in our survey!

We, Erik Driessen, Tony Artino, and Lauren Maggio, are inviting authors, such as yourself, who published HPE research in 2016 to participate in this health professions education (HPE) study. Please read the below information sheet carefully and feel free to contact us if you have additional questions by using the email addresses below.

Study purpose:
This study seeks to determine the prevalence of questionable research practices (QRPs) in HPE to inform educators, practitioners, and journal editors. In doing so, we hope the HPE community might be better positioned to take evidence-informed action, should the results indicate a need for such action. This research project includes this administration of an online survey and the aggregate analysis of collected survey data. We aim to publish our findings in a peer-reviewed journal.

Expectations:
If you agree to participate, we will ask you to complete an online survey. The survey can be completed on a computer or mobile device. We estimate the survey will take approximately 12 minutes. The survey, includes 43 Likert-type questions that ask you to indicate the prevalence of QRPs in your research. The survey also includes 24 additional items about publication pressure and basic demographic information (e.g., professional degrees earned, gender). No preparation is necessary for the survey.

Risk:
There is no direct advantage for you in participating in this study. However, we believe it might provide a better understanding of QRPs and thus may have advantages for the field of HPE research. It is up to you to decide whether or not to participate in the study. Participation is voluntary. If you do participate in the study, you can always change your mind and stop participating at any time during the study, without giving a reason. If you decide to stop the study, any data collected will be deleted and not used for this study. Whether you participate or not, there are no negative consequences for you. You will not be paid for your participation in this study. A potential disadvantage of participation is the time (approximately 15 minutes) it will take to complete the survey.

Data storage:
Data collection will begin upon clicking the below link to the survey. Each time you reach the end of a survey page and elect to continue, responses from that survey page will be submitted and stored. No personal identifiable information will be collected (i.e., we will not collect your name, email, or IP address). Data will be stored anonymously in Maastricht University's Qualtrics account and is not traceable to you. Also, your responses will be analyzed in combination with those of other respondents. Only Drs. Driessen, Artino and Maggio will have access to the anonymous data. All data will be reported in aggregate in any reports or publications.

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

If you participate in this study, you consent to the data being stored for 10 years after ending the study for further analysis within context of this project. You cannot participate in this study if you do not give permission for this data storage (by consenting below). After 10 years, the data will be destroyed.

This study has been approved by the Ethical Review Board Committee of the Netherlands Association for Medical Education (Dossier #937)

Please don't hesitate to contact us directly if you have any questions about the research or this survey or would like further information. We can be contacted at the following email addresses and phone numbers: Erik (e.driessen@maastrichtuniversity.nl; phone: 31(0)43-3885774)

We appreciate your time and expertise!
Erik, Tony and Lauren

If you consent to taking this study, please read and check yes below:

I have read the above information sheet for participants. I have had the opportunity to contact the investigators ask additional questions. My questions have been sufficiently answered. I have had enough time to decide whether to participate or not.

I know that participation is entirely voluntary. I am aware of my right to withdraw or end my participation from the study at any time. I do not need to justify that decision.

I know that certain people have access to my data. These people are listed in this information sheet. I am entitled to inquire and look into how my data are stored.

I consent to my data being used in the way and for the purpose stated in the information sheet. If for any reason my data would be used for research with another objective, I will be informed and again be asked to consent.

I consent to my data being stored for another 10 years after ending this study to permit further analysis within the context of this study.

I consent to participate in this study.

- O Yes
- O No

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/252982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Research Practices: In your work as an HPE researcher, **how often have you engaged in any of the following behaviors**, even if it has been only on a single occasion? If applicable, please consider your experiences with both quantitative and qualitative research.

	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable to my work
Conducted a human-subjects research study without ethics approval (i.e., without institutional review board [IRB] approval)	O	O	O	O	O	O	O
Circumvented one or more aspects of human-subjects ethics rules (i.e., IRB rules)	O	O	O	O	O	O	O
Collected course or curriculum data under the guise of "program evaluation" without human-subjects ethics (IRB) approval with the ultimate intent of using the data for research purposes	O	O	O	O	O	O	O
Inappropriately stored sensitive research data (e.g., data that contains personally identifiable information)	O	O	O	O	O	O	O
Inappropriately emailed sensitive research data (e.g., data that contains personally identifiable information)	O	O	O	O	O	O	O
Stopped collecting data earlier than planned because the results already reached statistical significance, without formal stopping rules	O	O	O	O	O	O	O
Fabricated data	O	O	O	O	O	O	O
Pressured a student or other subordinate to be a study participant in your research	O	O	O	O	O	O	O
Used students or residents as research subjects without informing the overseeing dean, program director, or other pertinent official	O	O	O	O	O	O	O

bioRxiv preprint first posted online Jan 31, 2018; doi:http://dx.doi.org/10.1101/252982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable to my work
Deleted data before performing data analysis without disclosure	O	O	O	O	O	O	O
Ignored a colleague's use of flawed data	O	O	O	O	O	O	O
Ignored a colleague's questionable interpretation of data	O	O	O	O	O	O	O
Reported a downwardly rounded p-value (e.g., reporting that a p-value of .054 is less than .05)	O	O	O	O	O	O	O
Misrepresented a participant's words or writings	O	O	O	O	O	O	O
Decided whether to exclude non-outlier data after looking at the impact of doing so on the results	O	O	O	O	O	O	O
In a qualitative study, failed to report disconfirming examples or cases that weaken your conclusions	O	O	O	O	O	O	O
Collected more data after seeing that the results were almost statistically significant	O	O	O	O	O	O	O
To confirm a hypothesis, selectively deleted or changed data after performing data analysis	O	O	O	O	O	O	O
Reported an unexpected finding as having been hypothesized from the start	O	O	O	O	O	O	O
Concealed results that contradicted your previous findings or convictions	O	O	O	O	O	O	O

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable to my work
Claimed you used a particular qualitative research approach appropriately (e.g., grounded theory) when you knowingly did not	O	O	O	O	O	O	O
Claimed you used a particular qualitative research technique appropriately (e.g., saturation, triangulation) when you knowingly did not	O	O	O	O	O	O	O
Spread study results over more papers than is appropriate (so-called “salami slicing”)	O	O	O	O	O	O	O
Deliberately failed to mention important limitations of a study in the published paper	O	O	O	O	O	O	O
Deliberately failed to mention an organization that funded your research in the published paper	O	O	O	O	O	O	O
Inappropriately modified the results of a study due to pressure from a research advisor or other collaborator	O	O	O	O	O	O	O
Inappropriately modified the results of a study due to pressure from a funding agency	O	O	O	O	O	O	O
Failed to disclose relevant financial or intellectual conflicts of interest	O	O	O	O	O	O	O
Used someone else’s ideas without their permission or proper citation	O	O	O	O	O	O	O
Used sections of text from another author’s copyrighted material without permission or proper citation	O	O	O	O	O	O	O
Used sections of text from your own publications without proper citation (so-called “self-plagiarism”)	O	O	O	O	O	O	O

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

	Never	Once	Occasionally	Sometimes	Frequently	Almost always	Not applicable to my work
Selectively cited certain papers just to please editors or reviewers	O	O	O	O	O	O	O
Cited articles and or materials that you have not read	O	O	O	O	O	O	O
Selectively cited your own work just to improve your citation metrics	O	O	O	O	O	O	O
Reused previously published data without disclosure (co-called “duplicate publication”)	O	O	O	O	O	O	O
Used confidential information obtained as a reviewer or editor for your own research or publications	O	O	O	O	O	O	O
Refused to share data with legitimate colleagues	O	O	O	O	O	O	O
Added one or more authors to a paper who did not qualify for authorship (so-called “honorary authorship”)	O	O	O	O	O	O	O
Accepted authorship for which you did not qualify (so-called “honorary authorship”)	O	O	O	O	O	O	O
Demanded authorship for which you did not qualify (so-called “honorary authorship”)	O	O	O	O	O	O	O
Omitted a contributor who deserved authorship	O	O	O	O	O	O	O
Submitted (or re-submitted) a manuscript or grant application without consent from one or more of the author	O	O	O	O	O	O	O
Submitted the same manuscript to multiple journals at once (so-called “duplicate” or “double submission”)	O	O	O	O	O	O	O

bioRxiv preprint first posted online Jan 31, 2018; doi:http://dx.doi.org/10.1101/252982. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Publication Pressure: These items address publication pressure. Please indicate the extent to which you agree or disagree with the following statements, as they relate to your particular HPE context.

	Completely disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Completely agree
Without publication pressure, my scientific output would be of higher quality	O	O	O	O	O
My colleagues' assessments of me, based on my publications, are stressful	O	O	O	O	O
Publication pressure strains my relationships with fellow researchers	O	O	O	O	O
I suspect that publication pressure leads some colleagues (whether intentionally or not) to inappropriately manipulate their data	O	O	O	O	O
Publication pressure leads me to have serious doubts about the validity of HPE research results	O	O	O	O	O
In my opinion, the pressure to publish scientific articles has become too high	O	O	O	O	O
My colleagues judge me mainly on the basis of my publications	O	O	O	O	O
I cannot share innovative research proposals with my colleagues for fear of those ideas being stolen	O	O	O	O	O
Publication pressure harms science	O	O	O	O	O

bioRxiv preprint first posted online Jan 31, 2018; doi:http://doi.org/10.1101/252982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Demographics:

What is your gender?

O Male

O Female

What is your age? [Drop-down menu]

- 1 to 100

In which country or region do you primarily work? [Drop-down menu]

▼ Africa ... Other

- Africa
- Asia
- Australia/New Zealand
- Canada
- Caribbean
- Europe (not including the UK
- Middle East
- South/Latin America
- United Kingdom
- United States
- Other

What is your current academic rank or position title? (Select one)

▼ Medical Student (1) ... Not Applicable (14)

- Medical Student
- Resident
- Graduate/PhD Student
- Fellow
- Postdoc
- Instructor
- Lecturer
- Assistant Professor
- Associate Professor
- Professor
- Professor Emeritus
- Staff
- Other
- Not Applicable

---

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Which degree(s) do you hold? (Check all that apply)

- ☐ Bachelor's degree (BS, BA, BSN, etc.)
- ☐ Master's degree (MA, MS, MSW, MPH, MSN, etc.)
- ☐ Professional medical degree (MD, DO)
- ☐ Doctoral degree (PhD, EdD, DrPH, etc.)
- ☐ Other professional degrees (JD, PA, DVM, etc.)

For the highest degree you selected above, what is your primary area of study? (Select one)

- O Basic Science
- O Clinical Science
- O Social Science/Education
- O Humanities
- O Other: _____

What is your primary work role? (Select one)

- O Clinician
- O Administrator or Program Director
- O Teacher (clinical or classroom setting)
- O Researcher
- O Other: _____

---

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982. This version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Maastricht University

Which of the following best describes the context in which you work? (Check all that apply)

- ☐ Undergraduate Medical Education (UME)
- ☐ Graduate Medical Education (GME)
- ☐ Continuing Medical Education (CME)

In a typical work week, approximately what percentage of your work time do you spend on health professions or medical education research activities, including writing up your research (please report your answer as a percentage)?

_____

How many years have you been involved in health professions or medical education (in any capacity)?

_____

How many years have you been involved in conducting research in health professions or medical education?

_____

In thinking about your primary research activities, how would you characterize your work (Select one)?

- O I am a quantitative researcher
- O I am a qualitative researcher
- O I am a mixed-methods researcher

In total, approximately how many peer-reviewed publications have you authored or co-authored?

_____

---

bioRxiv preprint first posted online Jan 31, 2018; doi: http://doi.org/10.1101/256982; this version posted April 12, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

If you have any other comments related to questionable research practices in HPE, or comments about this questionnaire, please share those here:

_____

_____

_____

_____

_____

Note that by clicking the below "next" button you will submit your questionnaire responses and be unable to return to the questionnaire.

Next

We thank you for your time spent taking this survey.
Your response has been recorded.

Powered by Qualtrics