International Journal of Medical and Health Sciences Research

URL: www.pakinsight.com

STRATEGIES FOR ADDRESSING POOR STATISTICAL LITERACY AMONG HEALTH SCIENTISTS AND SCIENCE COMMUNICATORS

Raywat Deonandan[1]

[1]*Interdisciplinary School of Health Sciences, University of Ottawa, Ottawa, Canada*

ABSTRACT

Statistical literacy (SL) is an aspect of numeracy and a core educational competency. Overall levels of SL in North American society may be in decline, while evidence abounds of poor SL among researchers globally. Three strategies are recommended to help address poor SL among scholars and science communicators, with a particular focus on the health and medical sciences: (1) mandatory inclusion of computerized statistical platforms in undergraduate statistics classes; (2) requirement for statistical review in peer-reviewed paper submissions; and (3) the exploration of novel methods of results communication.

© 2016 Pak Publishing Group. All Rights Reserved.

Keywords: Statistics, Literacy, Research, Knowledge translation.

1. INTRODUCTION

There is much banter in the educational media about shifting levels of literacies across several domains. Recent furor has been spurred by the publishing of results from the Organization for Economic Cooperation and Development's (OECD) 2012 study, Programme for International Student Assessment [1, 2] which sought to measure science and mathematical literacy among residents of the OECD's 44 member nations. North Americans in particular did not fare well in the report, which found that a large number of Canadians can master complex problem solving skills, but are unable to work with numbers at a high level of competency. Indeed, 23 percent of Canadians performed at the lowest level of numerate competency, compared with the OECD average of 19 percent. Americans' math and problem-solving skills lagged even further behind those of their OECD counterparts.

An important dimension of numeracy is statistical literacy (SL), with basic numeracy being a prerequisite for becoming statistically literate. SL can be considered the ability to comprehend statistical language and meaning, when presented in a casual manner, as is relevant when a consumer is presented with advertising claiming that "4 out of 5 dentists" recommend a particular brand of toothpaste. It can also be considered a base skills set for scholars seeking to assess and communicate evidence generated from their own research. In the latter respect, a sustained focus on the maintenance of a high level of statistical competency among researchers is essential for the growth and proper application of scientific evidence.

There is an abundance of evidence that incomplete SL can result in a skewed appreciation of both the messaging and the importance of research studies. An improper distinction between relative risk and absolute risk, for example, has been shown to be associated with a more enthusiastic acceptance of cancer screening as being scientifically worthwhile [3] a finding that may not be supported by the data being reviewed. Poor SL among consumers is problematic in that it impairs the ability to accept and act on important evidence arising from research. But poor SL

among other actors on the research spectrum (i.e., scientists and science communicators) is particularly worrisome, in that a poorly analyzed and poorly translated piece of research may inappropriately inform action and policy, regardless of how highly numerate the end user might be. In other words, unless the initial research is well analyzed and appropriately presented, its downstream effects upon society are likely to be deleterious, even if end-user consumers have adequate SL.

With this paper, I suggest three strategies for improving SL among health-based researchers and science communicators (often called 2nd and 3rd level users [4]: (1) The mandatory inclusion of the use of computers and popular statistical platforms in undergraduate statistics classes; (2) the requirement of dedicated statistical review in peer-reviewed paper submissions; and (3) the exploration of novel methods of results communication, by both researchers and science communicators.

2. STATISTICAL SOFTWARE

Like most things, software experiences trends in popularity and fashion. However, for the analysis of very large datasets, particularly large population surveys, SAS and its close relative SUDAAN are still the industry standard for population health studies, despite having been first introduced more than four decades ago. Meanwhile, SPSS continues to enjoy popularity among social scientists and undergraduate teachers, in part due both to its ease of use and affordable student version. While not relevant to a discussion specifically of statistics, qualitative analysis programs, like NVIVO, are gaining in popularity in both research labs and universities. Statistical consultants enjoy more targeted and nuanced statistical packages, such as S-PLUS and its open source cousin, R. The famous open source epidemiological package Epi-info has fallen out of fashion, but it is still available for free download, and offers much flexibility in both data collection and analysis options. Most commercial and open source spreadsheet packages offer sufficient statistical analysis options to satisfy many researchers seeking uncomplicated analytical solutions. One open access, free option, called PSPP, mimics the operability of SPSS and represents a solution for emerging, lower income economies struggling to contribute to the global scientific output.

As well, there exists a plethora of online data analysis solutions. The online options, while convenient, are a bit problematic. A Google search of "online sample size" results in 146,000 results, most seemingly of web-based sample size calculators. Most, like that of the Statistics department of the University of British Columbia [5] offer a degree of control of the parameters for the sample size computation, such as whether the test is one- or two-sided, and the desired power level. What is missing is a nuance of understanding. The assumption underlying all convenient online calculators is that the user understands the nature and limits of the options that he or she enters, whereas a manual computation often necessitates a full exploration by the researcher of the limitations and known qualities of his or her data. In other words, while web-based options are convenient, they should only be sought by those with a demonstrated confidence in the assumptions underlying the desired calculation.

Statistical packages like SAS and R, on the other hand, demand a degree of statistical fluency on the part of the user. A novice will struggle with the syntax of the programming languages (even with the point-and-click options of the GUI interface) and with information that must be entered, such as the formatting of the data. This struggle is useful from an educational standpoint. It serves as a filter for comprehension, such that a minimum understanding of statistical terminology and assumptions is necessary to use the system. We must remain mindful, however, that command-line fluency is often the domain of the rarefied specialist, while GUI platforms can sometimes assist in gaining understanding [6]. There appears to be consensus, however, that the use of computers as either an analytical tool, or for the provision of simulation exercises, is a teaching approach to be celebrated in introductory classes [7].

From a larger, realpolitik standpoint, the introduction of software early in the undergraduate education system serves to prepare scholars for the genuine environment of a professional research existence, in which a reliance on statistical software is the norm. There is already some evidence that computer usage in this fashion may improve overall SL [8]. While the manual computation of simple descriptive and bivariate relationships is useful for helping

to ingrain an appreciation for the mechanics of frequentist tests being applied, ultimately it is the use of data processing software that ushers a student into understanding how statistics shapes his or her modern world.

In terms of career planning, an early introduction to software eases scholars into the worlds of so-called "big data", "meta data" and limitedly into Bayesian methods, which all seem to be indicative of the statistics paradigm of the immediate future, and which frankly are altogether inaccessible (except, possibly, for Bayesian methods) without the use of computers.

3. STATISTICAL REVIEW

Poor statistical reporting in science journals, particularly health science journals, is at crisis levels. The trend of misreporting findings or exaggerating small effects is now sufficiently noticeable that it is not an uncommon topic in lay publications, as in a 2010 article in The Atlantic, which stated that, "Much of what medical researchers conclude in their studies is misleading, exaggerated, or flat-out wrong." [9]. This is a global phenomenon, with one study suggested that almost 78% of the original articles published in two Indian pharmacology journals had used inappropriate statistics [10].

Presently, a percentage of peer-reviewed journals institute a specific review of submitted statistical content. While in some countries, and in some disciplines, that percentage can be as high as 80% [11] most journals roll statistical review into a general methodological review that is usually part and parcel of the grander overall paper review. The main reason for this is the lack of willing peer-reviewers, and the ever diminishing time available for busy scholars to offer review services for free. There are a few notable exceptions, particularly among the highest impact journals. Among those exceptions is *The Lancet*, which maintains a specific list of statistical and methodological reviewers, one or more of whom are attached to any submitted study with an important statistical component. Journals of *The Lancet*'s stature can attract high quality reviewers. But journals with less fame and perceived quality may struggle in this regard.

However, it is useful for at least an effort to be made to communicate that a journal will give special attention to submitted papers' statistical content. It sends a message that SL is a critical component to scholarship that deserves special attention. This is doubly important as the peer-review process overall is coming under scrutiny, in the wake of revealing studies indicating professional misconduct and scientific fraud to be more prevalent than previously assumed [12]. Methodological rigour and SL are useful prophylaxes against much misconduct, as they constitute objective tests to be applied to research claims.

Editors in and of themselves do not have sufficient training and expertise to review studies' statistical content. The recruitment of statistical and methodological specialists from the community of scholars is an important component to the role of dedicated statistical review in encouraging overall SL. Mladen Petrovečki suggests that the role of a dedicated statistical reviewer is twofold: to ensure that the study was properly conducted and results appropriately presented and to reveal possible errors and omissions [13]. The different types of errors are summarized by Altman [14] into 4 categories: errors in study design (e.g., no randomization in controlled trials or an inappropriate control group), errors in data analysis (e.g., using a given statistical test on a type of data for which it does not apply), errors in data presentation (e.g., using standard error instead of standard deviation to describe data breadth), and errors in data interpretation (e.g.,mistaking a correlation for a causal relationship).

In general, Petrovečki recommends that studies undergo three separate reviews: the general review which he calls a "professional" review, which includes relevance, sufficiency of novelty, proper citation, and all the other qualities of good scientific writing; ethical review, which can be satisfied through noting that the study received approval from a recognized Research Ethics Board; and lastly statistical review, which seeks to determine if the aforementioned errors have been committed, and whether the design of the study was appropriate to the research questions and analytical techniques applied.

Dedicated statistical review is not a panacea. Indeed, there is evidence to suggest that statistical and methodological reviews are often ignored and their recommended changes rarely manifest as improvements in the final manuscript [15, 16]. What may be needed to address this failing is stronger editorial resolve, such that greater editorial decision-making is placed on the opinion of the statistical or methodological reviewer. Assuming such editorial resolve, the implementation of a dedicated statistical review accomplishes two important goals. First, it helps assure that only the best evidence is deemed appropriate for entry into the vetted public discourse. And second, it sets a high standard for statistical excellence that all researchers must seek to attain.

4. EXPLORATION OF STATISTICAL COMMUNICATION TOOLS

Perhaps the greatest challenge in seeking global SL is the inability to effectively communicate statistical results without sinking into the mire of dry, technical language that ends up disenfranchising the relevant audience and obscuring any important nuance of the results. It can also open the door to misinterpretation of those results. This can happen when the complicated and detailed content of an important piece of scientific research is simplified for and by the lay media, often to the detriment of subtlety and to the exclusion of important caveats.

It is therefore incumbent upon science communicators --a group that can include journalists, teachers, and scientists themselves-- to use a diverse set of communication tools to more properly and effectively convey the truth and essence of research results. In recognition of the importance of this skill, the UK Royal Statistical Society offers an award for statistics in journalism [17] meant both to reward proper use and communication of statistics in the law media, and to encourage the improvement of such communication among the cadre of media professionals.

The set of tools available to communicators is wide indeed, and each brings with it opportunities and caveats. There is evidence that visual tools are more effective in conveying statistical nuance than is simply the stating of probabilities [18] whereas natural frequencies are best for reaching those with exceptionally low numeracy [19]. Similarly, the use of new media, such as text messages, to convey statistical results to a lay audience found variable effectiveness depending upon the recipients' levels of math anxiety [20].

Obviously, there is no single universal solution to the challenge of reaching an audience with variable levels of SL. Barriers are multifold. Indeed, larger social constructs, such as socioeconomic status and overall reading literacy levels appear to have a role in gating SL [21]. However, we live in an age of communications miracles, with rich media, in the form of streaming video, audio, hypertext and static visual portrayals, available to consumers at the touch of a keyboard. Reliance upon plain text, typically expressed through passionless technical language, is an outdated approach that offends the complexity of our media-driven society.

5. THE ROLE OF WORKSHOPS

Related to the wealth of rich media tools is the role of the larger scholarly community in ensuring high quality statistical content. Increasingly, the lack of mathematical content is seen as a marketing advantage when promoting educational instruments, such as accounting classes [22] This tack is understandable given the growing culture of fear surrounding many aspects of quantitative analysis in paedagogy, as in the title of Song & Slate's paper, "Taking Statistics Doesn't have to be Scary: Keeping the Heartrate Down." [23]. In it, the authors argue that there exist strategies for minimizing the anxiety that often accompanies the learning of statistics. Their strategies have high concordance with those described in the present paper, but include an additional stress on interpersonal components.

Along with reducing overall fear and anxiety surrounding statistics, the embracing of less technical and more accessible language has emerged as a tactic celebrated by such international bodies as the United Nations' National Statistics Office [24]. The Office also advocates for the production of a regular journal dedicated to the communication of important statistics using simple language. Most relevant to this paper is the Office's championing of seminars and public meetings for the purpose of explaining to a lay audience the nuances of their analyses.

It seems reasonable that one should encourage emerging scholars to maintain a continuous relationships with their larger professional communities for the purposes of both networking and skills development. With respect to statistical literacy, the plethora of workshops and seminars available both in-person and online provide an opportunity for convenient and affordable statistical skills enhancement. Such workshops can be integrated into existing professional development events that are already well attended by active members of many professions.

6. CONCLUSION

One way to think about SL is as a core competency: the ability to read and interpret data presented as evidence for arguments [25]. In this respect, it is a vital skill for complete participation in modern society, and a platform for broader critical thinking. Yet some studies suggest that medical professionals are dramatically under-skilled with respect to basic understanding of statistics, while simultaneously perceiving their statistics education to be adequate [26]. With our civilization accelerating its reliance upon technologies and the evidence needed to manage those technologies, the need to recognize a crisis in SL among both scholars, communicators, and policy makers is dire indeed.

Strategies are available to us to help address the gap in SL. But they must be implemented at the early stages of both the educational and research processes. With this paper, I have proposed three specific approaches. But as part of a wider campaign to re-invigour core literacy competencies in our education system, these strategies need to be partnered with an overall ethic of skills enhancement relevant to producing an intellectually well-equipped and evidence-savvy society.

REFERENCES

[1]   Organization for Economic Cooperation and Development, "Program for international student assessment: 2012 Results." Available http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm, 2012.

[2]   CBC News, "OECD study reveals Canada's polarized workforce." Available http://www.cbc.ca/news/business/oecd-study-reveals-canada-s-polarized-workforce-1.1930222, 2013.

[3]   T. J. Caverly, A. V. Prochazka, I. A. Binswanger, J. S. Kutner, and D. D. Matlock, "Confusing relative risk with absolute risk is associated with more enthusiastic beliefs about the value of cancer screening," *Med Decis Making,* vol. 34, pp. 686-692, 2014.

[4]   P. Kovacs, "Issues and challenges of statistical literacy in higher education," presented at the 59th ISI World Statistics Congress, 2013.

[5]   R. Brant, "Web based sample size/power calculations. University of British Columbia, Department of Statistics." Available: http://stat.ubc.ca/~rollin/stats/ssize/, 2014.

[6]   A. Unwin and H. Hofmann, "GUI and command-line –conflict or synergy? In K. Berk Pourahmadi, M. (Ed.). Computing science and statistics," in *Proceedings of the 31st Symposium on the Interface*, 1999.

[7]   T. S. Porter and T. M. Riley, "The effectiveness of computer exercises in introductory statistics," *Journal of Economic Education,* vol. 27, pp. 291-299, 1996.

[8]   A. Bakker, "Do computer tools improve statistical literacy?," *MSOR Connections,* vol. 5, pp. 28-29, 2005.

[9]   D. H. Freedman, "Lies, damned lies, and medical science." Available http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269, 2010.

[10]   Y. P. Jaykaran, "Quality of reporting statistics in two Indian pharmacology journals," *J. Pharmacol Pharmacother,* vol. 2, pp. 85–89, 2011.

[11]   S. Kannan, S. P. Deshpande, N. J. Gogtay, and U. M. Thatte, "Policy of reviewing statistics in Indian medical and surgical journals," *J. Pharmacol Pharmacother,* vol. 4, pp. 139–140, 2013.

[12]   F. C. Fang, R. G. Steen, and A. Casadevall, "Misconduct accounts for the majority of retracted scientific publications," in *Proceedings of the National Academy of Sciences of the United States of America*, 2012.

[13]   M. Petrovečki, "The role of statistical reviewer in biomedical scientific journal," *Biochemia Medica.,* vol. 19, pp. 223-230, 2009.

[14]   D. G. Altman, "Statistical reviewing for medical journals," *Stat. Med.,* vol. 17, pp. 2661-2674, 1998.

[15]   D. L. Schriger, R. J. Cooper, R. L. Wears, and J. F. Waeckerle, "The effect of dedicated methodology and statistical review on published manuscript quality," *Annals of Emergency Medicine,* vol. 40, pp. 334-337, 2002.

[16]   F. C. Day, D. L. Schriger, C. Todd, and R. L. Wears, "The use of dedicated methodology and statistical reviewers for peer review: A content analysis of comments to authors made by methodology and regular reviewers," *Annals of Emergency Medicine,* vol. 40, pp. 329-333, 2002.

[17]   Royal Statistical Society, "Awards for statistical excellence in journalism." Available: http://goo.gl/Yx0sWD, 2014.

[18]   R. Garcia-Retamero and U. Hoffrage, "Visual representation of statistical information improves diagnostic inferences in doctors and their patients," *Soc. Sci. Med.,* vol. 83, pp. 27-33, 2013.

[19]   M. Galesic, G. Gigerenzer, and N. Straubinger, "Natural frequencies help older adults and people with low numeracy to evaluate medical screening tests," *Med. Decis. Making,* vol. 3, pp. 368-371, 2009.

[20]   K. J. Silk and R. L. Parrott, "Math anxiety and exposure to statistics in messages about genetically modified foods: Effects of numeracy, math self-efficacy, and form of presentation " *J. Health Commun.,* vol. 19, pp. 838-852, 2014.

[21]   P. G. Doyle, "Developing statistical literacy with students and teachers in the secondary mathematics classroom," Master's Thesis, University of Waikato, 2008.

[22]   G. Mostyn, "Math in accounting? Not as much as you think. Worthy & James Publishing." Available: http://www.worthyjames.com/office-t-0908.html, 2006.

[23]   H. Song and J. Slate, "Taking statistics doesn't have to be scary: Keeping the heartrate down. University of South Carolina Aiken." Available http://www.usca.edu/essays/vol172006/slate.pdf, 2006.

[24]   United Nations, "Promoting statistical literacy. National statistics office." Available http://unstats.un.org/unsd/dnss/, 2015.

[25]   M. Schield, "Statistical literacy: Thinking critically about statistics. Association of public data users, of significance." Available http://web.augsburg.edu/~schield/milopapers/984statisticalliteracy6.pdf, 1999.

[26]   B. L. Anderson, S. Williams, and J. Schulkin, "Statistical literacy of obstetrics-gynecology residents," *Journal of Graduate Medical Education,* vol. 2, pp. 272-275, 2013.