Title: Being honest with causal language in writing for publication

Submit to: Editorial, *Journal of Advanced Nursing*

Authors: Deependra K **THAPA***, Denis C **VISENTIN*****, Glenn E **HUNT,** Roger **WATSON,** Michelle **CLEARY**,

Deependra K. **THAPA,** MPH, MSc, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. *Corresponding author: Deependra Kaji Thapa, College of Health and Medicine, University of Tasmania, Locked Bag 5052, Alexandria, NSW 2015, Australia. Email: deependrakaji.thapa@utas.edu.au ORCID: https://orcid.org/0000-0002-5689-0837

Denis C VISENTIN, PhD, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. Email: denis.visentin@utas.edu.au ORCID: https://orcid.org/0000-0001-9961-4384

Glenn E. HUNT, PhD, Discipline of Psychiatry, Concord Clinical School, University of Sydney, NSW, Australia. Email: glenn.hunt@sydney.edu.au ORCID: http://orchid.org/0000-0002-8088-9406

Roger WATSON RN, PhD, FAAN, Faculty of Health Sciences, University of Hull, Hull. UK email: r.watson@hull.ac.uk ORCID: https://orcid.org/0000-0001-8040-7625

Michelle CLEARY, RN, PhD, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. Email: michelle.cleary@utas.edu.au ORCID: http://orcid.org/0000-0002-1453-4850

Conflict of Interest statement: Nil

Author contributions: All authors have agreed on the final version and meet at least one of the following criteria recommended by the ICMJE (http://www.icmje.org/recommendations/)]:

- Substantial contributions to conception and design, acquisition of data or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content.

Being casual with causal language in writing for publication

Introduction

The misleading use of causal language in publication is problematic for authors, reviewers, and consumers of the information. Published research in quality journals has important knowledge implications and it is, therefore, contingent on authors to use language that is accurate and appropriate to their work. Language implying causal relationships, when this is not supported, may overstate the evidence-base especially if accepted by uncritical readers or unwitting members of the general public who may not understand how to interpret inferential statistics. The choice of language shapes academic thought across a range of areas including research, policy, clinical practice, and education impacting current and future directions within the research field (Kueffer & Larson, 2014). In an age when, with the aim of selling copy, the media vacillates on the benefits of a range of 'fads' around diet, alcohol and exercise, it can be very confusing for consumers of health services when, one day, what purports to kill you, thereafter is lauded as beneficial to health. Likewise, in the field of complementary and alternative medicines—including traditional Chinese medicine (Watson & Xue 2019)—where the evidence-base is threadbare, spurious relationships—which are certainly not causal—are touted as evidence of cause and effect people ought to be protected from false claims.

Commented [RW1]: Watson R, Xue C (2019) Publishing studies in traditional Chinese medicine Journal: *Journal of Advanced Nursing* doi: 10.1111/jan.14297

1. The nature of causality

Identifying causal relationships is a core objective of research. Establishing causation permeates the philosophy of science (Mackie, 1980), and gives coherence and temporal order to our understanding of nature (Flaherty, 2011). Rothman (1995) defines a cause as: "*…an act* *or event or a state of nature which initiates or permits, alone or in conjunction with other causes, a sequence of events resulting in an effect*" (page 91). In terms of health outcomes, Bonita et al. (2006) identifies causal factors as events, conditions, characteristics which have a role in the health outcome, either alone or in combination. Both definitions identify causal relationships involving an exposure and a resulting outcome. A cause is 'sufficient' when its presence always produces or initiates the effect and is 'necessary' if the outcome does not occur in its absence (Bonita et al., 2006). An example of a causal relationship is developing tuberculosis. While tuberculosis bacterium is the necessary cause, there are other contributing factors such as compromised immunity and co-morbidities which determine whether tuberculosis arises in an individual. In health research, there is almost always more than one factor contributing to the outcome of interest, making identification of causality more nuanced.

Causal inference is the process of finding an association between two factors and determining whether this is likely to be causal. Establishing a causal relationship requires not only statistical evidence but also clinical knowledge. Once an association is established, non-causal explanations including bias, chance and confounding need to be considered (Rothman & Greenland, 2005). Hill (1965) provided a set of 'considerations for causations' which provide weight to a causal interpretation for the association, including; strength, consistency, specificity, temporality, biological gradient, plausibility, coherence, experimental evidence, and analogy.

2. Study design and causality

There is a consensus that 'temporality', which refers to the necessity that the cause must precede the effect in chronological time, is essential for establishing a causal relationship, for which longitudinal follow-up is required. Well-designed longitudinal trials can provide inferences about causation (Green et al., 2011), while observational studies merely identify associations. Observational studies cannot provide evidence for the direction of causality, i.e. whether exposure influences the outcome, or vice versa, or whether both were influenced independently by some other confounding factors.

The concept of the *hierarchy in the levels of evidence* in terms of validity about causation implies that well-designed randomized controlled trials (RCT) possess the highest level of evidence (Buhse et al., 2018). Systematic reviews and meta-analyses provide better evidence than RCTs (Evans, 2003). Systematic reviews evaluate the quality of studies and the consistency of the results between studies, while meta-analyses increase the precision of estimates by analysing the pooled data from individual studies. Among observational studies research evidence level increases from, case reports, case series, cross-sectional, case-control, and cohort studies. The quality of methodology and design of the study also needs to be considered. A poorly designed and conducted RCT may provide biased evidence in contrast to a well-designed cohort study (Wang & Attia, 2010). As research designs differ in terms of the risk of error and bias of their results, grading evidence is important to assess the strength of causal evidence and if the effect is worthy of reporting (Visentin et al., 2020).

3. Language of causation in academic writing

The dissemination of research is a crucial step from evidence generation to evidence consumption and decision-making of clinical practitioners, policy makers, and the general public. Use of inappropriate causal language can yield inaccurate, imprecise, distorted and overstated claims (Boutron et al., 2014). Improper use of causal language such as misinterpreting correlations as causations can have potentially serious consequences, including erroneous medical decisions (Buhse et al., 2018) and contribute to misperceptions among potentially at risk groups (Richardson et al., 2014). Describing causal relationships in academic publications plays a crucial role in how we communicate scientific research within our community and to the public (Kleinberg & Hripcsak, 2011). Several reviews demonstrate that some published studies use problematic causal language (Glasziou & Chalmers, 2018; von Elm & Egger, 2004).

Causal language involves a clause or phrase where one event, state, action, or entity is explicitly presented as influencing another. There are several properties of linguistic expressions relating to causation. The use of a causative verb (e.g. increase, decrease, improve) actively in a sentence often portrays a strong causal relationship (Adams et al., 2017). Conjunctions such as 'because', 'due to', and 'since' are also used to express causal relations (Waldmann et al., 2017). Academic writers want to make their studies accessible, and interesting, which may lead them to use language that exaggerates the causality (Woloshin et al., 2009). Improper use of causal language may also be inadvertently used when the authors try to vary terms throughout to minimise monotony. For example, 'social support is associated with mental well-being of older adults' becomes 'social support improves mental well-being of older adults'. These two sentences may look similar but have different causal meaning. Writers may use cues such as can, could, may, appear to, before a verb for weaker expression or to indicate doubt (Adams et al., 2017), such as 'social support may improve mental well-being of older adults'. The choice of causal expressions could also be influenced by authors' first language (Kranich, 2011) in part due to differing causal linguistic expressions (Dunietz et al., 2017).

Using causal language in reporting the correlational findings from observational studies is a common source of misinformation (Boutron & Ravaud, 2018; Chiu et al., 2017). Haber et al. (2018) found 34% of studies used language too strong for strength of causal inference, with omitting confounding variables and generalizability being the most severe issues in causal inference. Rubin and Parrish (2007) found 70% of studies had methodological issues limiting causal inference, among which 60% used language which inflated the causal evidence. Such methodological reviews (Cofield et al., 2010; Glasziou & Chalmers, 2018; von Elm & Egger, 2004) typically classify causal language issues as binary; with causation and correlation the only possibilities. Sumner et al. (2014) provide seven levels of certainty in reporting finding: no mentioned relationship, statement of no relationship, statement of correlation, ambiguous statement of relationship, conditional statement of causation, statement of can, and statement of causation.

Causal language should indicate how one variable affects another including the direction of causality, while non-causal language describes a relationship, irrespective of the sequence of events. Choosing appropriate language is an important part of reporting the results of a study and their causal relationship. Table 1 provides a list of terminology of causation examples to describe causal and non-causal relationships.

Table 1. Terminology of causation*

Language implying causation	**Non-causal language**
• causes • effects, modifies	• associated • related • correlated

• increases/decreases • elevates/reduces • makes • improves • influences • impacts • results in • induces • effective in • is attributable to/contributes to • leads to • responsible for	• predicts • higher • lower • linked to • varies with

* Adapted from (Adams et al., 2017; Cofield et al., 2010; Zweig & DeVoto, 2015)

New research builds or extends upon findings from previous research. When using causal (or non-causal) language, authors should discuss how their research fits into the existing knowledge base as well the possible sources of bias considering the study design, and possible errors. Including a disclaimer and/or statement of study limitations may address the risk of readers misinterpreting findings. Cofield et al. (2010) recommended being explicit in using causal language especially in titles and abstracts because these are prominently displayed. Reviewers and editors should also consider rigorous evaluation of submitted manuscripts for the use of misleading language. Readers should carefully examine the study designs and sources of bias, and whether the results are correctly reported based on assumptions of the statistical tests used (Visentin & Hunt, 2017).

4. Conclusion

Scientific writing is different from general writing and language is crucial. Writing for publication, by clearly communicating findings while engaging the readership, can be challenging. Given the extent of exaggerated and misleading causal language, careful attention to the language of causation is required to minimise misinterpretation, and support knowledge development. Care in being casual with causal language ensures congruence with the study design and limits inflation of evidence.

References

Adams, R. C., Sumner, P., Vivian-Griffiths, S., Barrington, A., Williams, A., Boivin, J., . . . Bott, L. (2017). How readers understand causal and correlational expressions used in news headlines. *Journal of Experimental Psychology: Applied, 23*(1), 1-14. doi:https://doi.org/10.1037/xap0000100

Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic Epidemiology, 2nd edition* China: World Health Organization.

Boutron, I., Altman, D. G., Hopewell, S., Vera-Badillo, F., Tannock, I., & Ravaud, P. (2014). Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. *Journal of Clinical Oncology*, *32*(36), 4120-4126. doi:https://doi.org/10.1200/JCO.2014.56.7503

Boutron, I., & Ravaud, P. (2018). Misrepresentation and distortion of research in biomedical literature. *Proceedings of the National Academy of Sciences*, *115*(11), 2613-2619. doi:https://doi.org/10.1073/pnas.1710755115

Buhse, S., Rahn, A. C., Bock, M., & Mühlhauser, I. (2018). Causal interpretation of correlational studies – Analysis of medical news on the website of the official journal for German physicians. *PLOS ONE*, *13*(5), e0196833. doi:https://doi.org/10.1371/journal.pone.0196833

Chiu, K., Grundy, Q., & Bero, L. (2017). 'Spin' in published biomedical literature: A methodological systematic review. *PLoS Biology*, *15*(9), e2002173. doi:https://doi.org/10.1371/journal.pbio.2002173

Cofield, S. S., Corona, R. V., & Allison, D. B. (2010). Use of causal language in observational studies of obesity and nutrition. *Obesity Facts*, *3*(6), 353-356. doi:https://doi.org/10.1159/000322940

Dunietz, J., Levin, L., & Carbonell, J. (2017). Automatically tagging constructions of causation and their slot-fillers. *Transactions of the Association for Computational Linguistics*, *5*, 117-133. doi:https://doi.org/10.1162/tacl_a_00050

Evans, D. (2003). Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing*, *12*(1), 77-84. doi:https://doi.org/10.1046/j.1365-2702.2003.00662.x

Flaherty, M. G. (2011). *The textures of time: Agency and temporal experience*. Philadelphia: Temple University Press.

Glasziou, P., & Chalmers, I. (2018). Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers. *British Medical Journal*, *363*, k4645. doi:https://doi.org/10.1136/bmj.k4645

Green, M. D., Freedman, D. M., & Gordis, L. (2011). Reference guide on epidemiology. In National Research Council (Ed.), *Reference manual on scientific evidence: Third edition* (pp. 549-632). Washington DC: The National Academies Press. Retrieved from https://www.nap.edu/catalog/13163/reference-manual-on-scientific-evidence-third-edition. doi:https://doi.org/10.17226/13163

Haber, N., Smith, E. R., Moscoe, E., Andrews, K., Audy, R., Bell, W., . . . on behalf of the, C. r. t. (2018). Causal language and strength of inference in academic and media articles shared in social media (CLAIMS): A systematic review. *PLOS ONE*, *13*(5), e0196346. doi:https://doi.org/10.1371/journal.pone.0196346

Hill, A. B. (1965). The environment and disease: association or causation? *Journal of the Royal Society of Medicine*, *58*, 295-300. doi:https://doi.org/10.1177/0141076814562718

Kleinberg, S., & Hripcsak, G. (2011). A review of causal inference for biomedical informatics. *Journal of Biomedical Informatics*, *44*(6), 1102-1112. doi:https://doi.org/10.1016/j.jbi.2011.07.001
Kranich, S. (2011). To hedge or not to hedge: The use of epistemic modal expressions in popular science in English texts, English-German translations, and German original texts. *Text and Talk*, *31*(1), 77-99. doi:https://doi.org/10.1515/TEXT.2011.004
Kueffer, C., & Larson, B. M. H. (2014). Responsible use of language in scientific writing and science communication. *BioScience*, *64*(8), 719-724. doi:https://doi.org/10.1093/biosci/biu084
Mackie, J. L. (1980). *The cement of the universe : a study of causation*. Oxford: Clarendon Press.
Richardson, S. S., Daniels, C. R., Gillman, M. W., Golden, J., Kukla, R., Kuzawa, C., & Rich-Edwards, J. (2014). Society: Don't blame the mothers. *Nature*, *512*(7513), 131-132. doi:https://doi.org/10.1038/512131a
Rothman, K. J. (1995). Causes. 1976. *American Journal of Epidemiology*, *141*(2), 90-95. doi:https://doi.org/10.1093/oxfordjournals.aje.a117417
Rothman, K. J., & Greenland, S. (2005). Causation and Causal Inference in Epidemiology. *American Journal of Public Health*, *95*(S1), S144-S150. doi:https://doi.org/10.2105/ajph.2004.059204
Rubin, A., & Parrish, D. (2007). Problematic phrases in the conclusions of published outcome studies: Implications for evidence-based practice. *Research on Social Work Practice*, *17*(3), 334-347. doi:https://doi.org/10.1177/1049731506293726
Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., . . . Chambers, C. D. (2014). The association between exaggeration in health related science news and academic press releases: retrospective observational study. *British Medical Journal*, *349*, g7015. doi:https://doi.org/10.1136/bmj.g7015
Visentin, D. C., Cleary, M., & Hunt, G. E. (2020). The earnestness of being important: Reporting non-significant statistical results. *Journal of Advanced Nursing*, *(In press)*. doi:https://doi.org/10.1111/jan.14283
Visentin, D. C., & Hunt, G. E. (2017). What do the stats mean? Improving reporting of quantitative nursing research. *International Journal of Mental Health Nursing*, *26*(4), 311-313. doi:https://doi.org/10.1111/inm.12352
von Elm, E., & Egger, M. (2004). The scandal of poor epidemiological research. *British Medical Journal*, *329*(7471), 868–869. doi:https://doi.org/10.1136/bmj.329.7471.868
Waldmann, M. R., Solstad, T., & Bott, O. (2017). Causality and causal reasoning in natural language. In M. R. Waldmann (Ed.), *The oxford handbook of causal reasoning*. Oxford: Oxford University Press.
Wang, J. J., & Attia, J. (2010). Study designs in epidemiology and levels of evidence. *American Journal of Ophthalmology*, *149*(3), 367-370. doi:https://doi.org/10.1016/j.ajo.2009.08.001
Watson R, Xue C (2019) Publishing studies in traditional Chinese medicine Journal: *Journal of Advanced Nursing* doi: 10.1111/jan.14297
Woloshin, S., Schwartz, L. M., Casella, S. L., Kennedy, A. T., & Larson, R. J. (2009). Press releases by academic medical centers: Not so academic? *Annals of Internal Medicine*, *150*(9), 613-618. doi:https://doi.org/10.7326/0003-4819-150-9-200905050-00007
Zweig, M., & DeVoto, E. (2015). Observational studies: Does the language fit the evidence? Association vs. causation. Retrieved from https://www.healthnewsreview.org/toolkit/tips-for-understanding-studies/does-the-language-fit-the-evidence-association-versus-causation/