

Use of Causal Language in Observational Studies of Obesity and Nutrition

Stacey S. Cofield^a Rachel V. Corona^b David B. Allison^{a,c}

^a Department of Biostatistics, School of Public Health,

^b Department of Epidemiology, School of Public Health,

^c Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, USA

Keywords

Epidemiology · Nutrition · Obesity · Causal language · Observational studies

Summary

Objective: To assess the inappropriate use of causal language in studies on obesity and nutrition. **Methods:** Titles and abstracts of 525 peer-reviewed papers in the 4 leading journals in the fields of obesity and nutrition were scrutinized for language implying causality in observational studies published in 2006. **Results:** Such misleading language appeared in 161 papers (31%) independent of funding source. Remarkably 49% of studies lacking statistically significant primary outcomes used misleading language compared to 29% of those with p values ≤ 0.05 (chi square $p < 0.001$). Exculpatory language was present in the body of the text in 19% of the 161 studies. **Conclusion:** We suggest that editors and reviewers evaluate submissions for misleading reporting.

Introduction

Randomized controlled experiments are generally recognized as allowing strong inferences about causation [1] although studies of mechanisms require different designs. In contrast, observational, nonrandomized studies (including those typically labeled ‘epidemiologic’), though capable of demonstrating associations and supporting inferences, do not generally justify statements about causation. Language implying cause and effect relationships is sometimes inappropriately used in reports of observational studies and has the potential to mislead some scientifically trained readers and more so mass

media reporters, clinicians and the general public with less scientific background. Regardless of cause, journal editors, reviewers and authors of obesity and nutrition research papers need to be vigilant about inappropriately used causal language in reports of observational studies. This is especially so regarding titles and abstracts, which often are the reader’s only exposure to study results.

Our search of PubMed for the keywords ‘nutrition’ and ‘obesity’ identified over 14,000 titles [2]. Of these, 14% had only a title (no abstract) and 7% only a title and abstract (no full paper available). Only 22% of the 79% with a full text option had free access. Hence, we focused on titles and abstracts to assess the frequency of inappropriate use of causal language in observational studies in the four leading nutrition and obesity journals, although we analyzed the full text. Because of concerns that industry funding leads to biased reporting of nutrition and obesity research [3], we assessed whether the frequency of the use of causal language differed by source of funding.

Material and Methods

Data Sources

All reports of non-experimental, human observational studies from the 2006 volumes of the *American Journal of Clinical Nutrition*, *Journal of Nutrition*, *Obesity* and the *International Journal of Obesity* were reviewed for use of causal language in titles or abstracts. According to the Institute for Scientific Information’s (ISI) 2007 impact factor scores, these are the four highest ranking non-proceedings journals in ISI’s category of ‘nutrition and dietetics’ that publish original empirical papers. All articles published in 2006 in each of these journals were evaluated to determine eligibility for inclusion and assessed for other factors (author affiliation, funding source, etc. see below). A total of 1,617 articles were retrieved from 2006 issues of *American Journal of Clinical Nutrition*, *Journal of Nutrition*, *Obesity* and the *International Journal of Obesity* (fig. 1a).

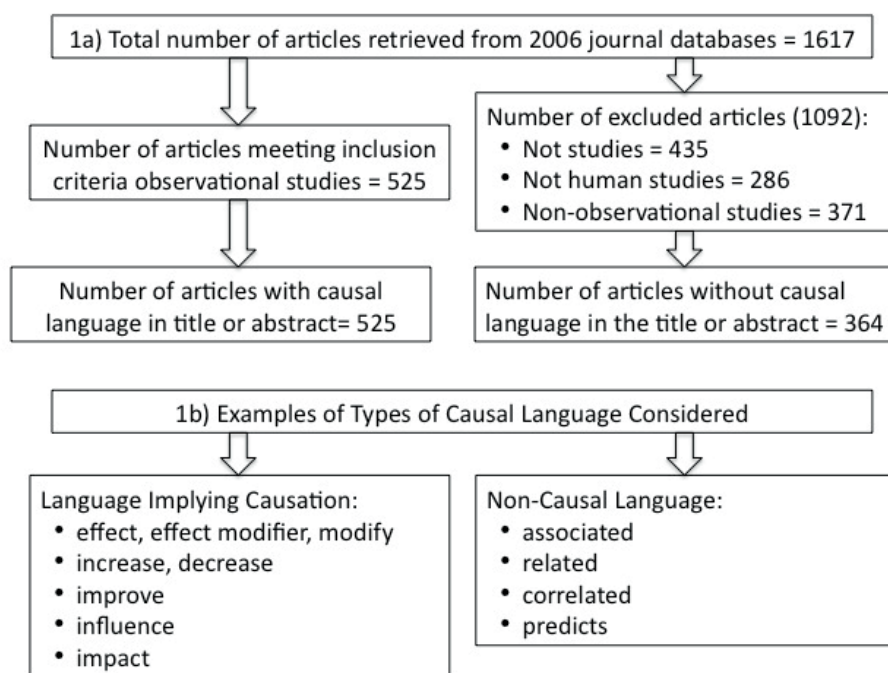


Fig. 1. a Protocol for article selection. **b** Examples of types of causal language.

Inclusion and Exclusion Criteria

We only included original human observational studies. Thus we excluded studies of non-humans (286 articles), methodological articles, reviews and meta-analyses, i.e. articles not describing empirical studies of raw data ($n = 435$) and experiments randomizing to an independent variable under investigation ($n = 371$). This resulted in 525 articles for inclusion (fig. 1a).

Determination of Causal Language

Each observational study was independently reviewed for the presence of language implying causation in the title and abstract by two reviewers from the UAB Departments of Biostatistics and Epidemiology. Causal language was classified as ‘none’ (no use) or ‘used’ in title, in abstract or in both title and abstract in which case the specific word or phrase indicating causality was recorded. Examples of language implying causation are shown in figure 1b. When such language was used, it was noted whether a statement disavowing causation was present in the Discussion section. Discrepancies between the independent reviewers were adjudicated by consensus of the 3 investigators.

Covariates

Other variables were recorded for use in secondary analyses: i) study design (cross-sectional, case control, or cohort), ii) p value for primary outcome of the study recorded as exact p value less than a given significance level (e.g. $p < 0.05$), and iii) source of funding as ‘non-industry funded’, ‘industry funded’ or ‘no listed funding source’ in Acknowledgements, Funding Source section or footnotes on the title page. To be classified as a non-industry funded study, the text had to specify a non-industry source, e.g. government agency, research university or non-profit independent research organization. Industry funded studies listed a for-profit affiliation. If a paper claimed: ‘There was no funding/outside support for this study’, but the author(s) were employed by industry, we considered the funding source to be industry funded. Studies by authors not directly affiliated with industry yet lacking a listed source of support were classified as having ‘no listed funding source’.

Statistical Analysis

Inter-rater reliability was assessed by percent agreement between 2 reviewers of papers assessed for any causal language use (title, abstract or

both) and by κ coefficients (\pm SE), a common metric of reliability for categorical data [4]. All descriptive statistics are reported as N and percent (%) of N. Proportion or causal language use was compared using a χ^2 test or Fisher’s Exact Test, where appropriate. Two-tailed p values are provided for descriptive purposes to indicate imbalances between groups, with α , the nominal type 1 error rate, set to 0.05. No adjustments were made for multiple testing. Data analysis was carried out using JMP version 8.0 (SAS Institute Inc; Cary, NC, USA).

Results

Inter-Rater Reliability

The reviewers agreed on use of any causal language, regardless of location (title, abstract, or both) in 512 of the 525 papers (97.5% agreement), resulting in a κ of 0.94 ± 0.016 . Eight of the 13 papers with reviewer disagreement (62%) contained casual language arrived at by consensus.

Descriptive and Inferential Statistics

Table 1 provides descriptive statistics of the presence of causal language according to journal, design, mention of statistical significance and funding source of the articles. Of the 525 papers, 161 (31%) used causal language in the title (7%), abstract (57%) or both (39%), and 31 (19%) included a statement that disavowed causation. The majority of the 525 articles were cross-sectional (266, 51%) whereas only 3% (18) were case control studies. There was nominally more causal language in case control studies compared to other study types (50% vs. 30% in cohort and 30% in cross-sectional studies; $p = 0.0660$, Fisher’s exact test). The majority of articles (419 or 80%) were non-industry funded; there was no significant difference in the use of causal language by source of funding ($p = 0.6825$).

Table 1. Use of causal language (Y/N) in leading nutrition journals 2006

		Number of articles (% of 525)	Number with causal language, (%)	p value*
Journal	AJCN	142 (27.0%)	23 (16.2%)	<0.0001
	IJO	174 (33.1%)	41 (23.6%)	
	JON	70 (13.3%)	39 (55.7%)	
	OBS	139 (26.5%)	58 (41.7%)	
Study design	case control	18 (3.4%)	9 (50.0%)	0.0660**
	cohort	241 (45.9%)	73 (30.3%)	
	cross sectional	266 (50.7%)	79 (29.7%)	
Manuscript primary result p-value	'not significant' (>0.05)	69 (14.0%)	34 (49.3%)	0.0006
	'significant' (≤0.05)	423 (86.0%)	121 (28.6%)	
Industry funding source	no	419 (79.8%)	125 (29.8%)	0.6774
	not indicated	64 (12.2%)	21 (32.8%)	
	yes	42 (8.0%)	15 (35.7%)	

AJCN = *American Journal of Clinical Nutrition*; JON= *Journal of Nutrition*; OBS = *Obesity*; IJO = *International Journal of Obesity*.

*Chi-square comparison of number with causal language.

** Fisher's exact test, $p < 0.05$ considered significant difference.

Only 6% Of the 525 papers did not include a p value associated with the primary endpoint; of those that did 86% (423) had p values < 0.05 . While 29% of papers with a significant p value used causal language, 49% of those with $p \geq 0.05$ used causal language ($p = 0.0006$). The location of causal language differed by the significance level associated with the primary endpoint ($p = 0.0016$). 12 (17%) of the papers with a nonsignificant p value used causal language in both the title and abstract while 32% (22) of these papers used causal language in the abstract only. For papers with a significant result associated with the primary outcome, 12% (50) used causal language in the abstract and title, 15% (65) in the abstract only and 1% (6) in the title only.

Discussion

The quality of reporting of observational studies is an important issue [5]. We found inappropriate use of causal language in the abstracts or titles of almost one third of human observational obesity or nutrition related study reports in four major obesity and nutrition journals. This is troubling because it misleads readers, particularly those who only peruse titles and abstracts. This may be especially true for less scientifically trained readers including reporters in the mass media who often have little scientific training and often describe themselves as having difficulty with key tasks involved in scientific reporting [6, 7]. Only a small proportion of articles included a disclaimer or statement of study limitations further increasing the risk of readers misapprehending the importance of the findings. While observational studies can contribute to the evidence supporting an accurate inference of causation, it re-

quires understanding of design, analysis and statistical interpretation to determine whether a single observational study was designed, analyzed and executed sufficiently well to warrant an inference of causation.

Notably, we found that source of funding was not a determinant of the inappropriate use of causal language. In this one aspect then, in contrast to implications in the literature [8], yet consistent with other recent findings [9], these results suggest that industry-funded research is not necessarily reported any less accurately than is non-industry-funded research. However, only a relatively small proportion of observational studies in this area were funded by industry, leaving us with only 42 such studies to use in this comparison, a limitation of our study which lacked the power to determine a statistically significant difference. There are two other limitations to our analyses. Whereas we examined the title and the abstract for use of causal language in only four journals, we did select the four highest ranked by ISI's 2007 impact factor score. Our analysis was limited to the fields of obesity and nutrition; results might be different in other fields.

Given the limitations, the implications of our findings are severalfold. First, readers should be aware that the fact that statements implying causation can be found in the abstracts of peer-reviewed papers or quoted in secondary sources citing those papers does not necessarily imply that conclusions of causation are warranted but that readers have to carefully examine study designs themselves. Second, the fact that the rates varied so much by journal suggests that editorial and/or authorship practices may vary substantially among journals, leaving some with more opportunity for improvement. Third, our results can be seen as setting a benchmark of practices in 2006. Having established a protocol, it should not be difficult

for editorial staff to collect such information on an ongoing basis as a way of assessing the rigor of their own editorial practices and monitoring for improvements over time.

Acknowledgments

Supported in part by the National Institutes of Health (P30DK056336). The opinions expressed are those of the authors and not necessarily those of the NIH or other author affiliated organizations.

The authors' responsibilities were as follows:

DBA: conception of project, development of protocol, drafting of manuscript, interpretation of information; SSC: oversight of protocol implementation, full access to and analysis of data, drafting of manuscript, interpretation of information; RVC: Retrieval of literature, drafting of manuscript, interpretation of information. DBA is the guarantor of the paper, having had full access to all of the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Disclosure

The authors' competing interests were as follows:

DBA has received grants, honoraria, donations, and consulting fees from numerous food, beverage, pharmaceutical companies, and other commercial, government, and nonprofit entities with interests in obesity and nutrition; DBA also serves or has served on the editorial boards or as an associate editor for three of the four journals examined herein and for competing journals; SSC has served as a reviewer for one of the journals and has received consulting fees from GSK and has clinical trial grants that receive contributions from Amgen, Biogen IDEC, and TEVA Neuroscience. RVC does not have any competing interests.

References

- 1 Green MD, Freedman DM, Gordis L: Reference Guide on Epidemiology. In Federal Judicial Center, Reference Manual on Scientific Evidence 2000. [www.fjc.gov/public/pdf.nsf/lookup/sciman06.pdf/\\$file/sciman06.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/sciman06.pdf/$file/sciman06.pdf).
- 2 PubMed: National Center for Biotechnology Information, U.S. National Library of Medicine. Retrieved June 9, 2010, from www.ncbi.nlm.nih.gov/pubmed.
- 3 Lesser LI, Ebbeling CB, Goozer M: Relationship between funding source and conclusion among nutrition-related scientific articles. *PLoS Med* 2007;4:e5.
- 4 Cohen J: A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
- 5 von Elm E, Altman DG, Egger M: The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007;4:e296.
- 6 Saari MA, Gibson C, Osler A: Endangered species: science writers in the Canadian daily press. *Public Underst Sci* 1998;7:61–81.
- 7 Voss M: Checking the pulse: Midwestern reporters' opinions on their ability to report health care news. *Am J Public Health* 2002;92:1158–1160.
- 8 Huff J: Industry influence on occupational and environmental public health. *Int J Occup Environ Health* 2007;13:107–117.
- 9 Thomas O, Thabane L, Douketis J: Industry funding and the reporting quality of large long-term weight loss trials. *Int J Obes* 2008;32:1531–1536.