



Make quotations great again: a proposal to reduce false-knowledge

Ronen Bareket ^{1,2} Mara A Schonberg,³
Yochai Schonmann ^{1,2,4}

10.1136/bmjebm-2018-111109

¹Department of Family Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

²Department of Family Medicine, Rabin Medical Center and Tel Aviv & Dan districts, Clalit Health Services, Tel Aviv, Israel

³Department of Medicine, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA, USA

⁴Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

Correspondence to:
Dr Ronen Bareket,
Departement of family medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv, 69978, Israel;
ronenba@gmail.com



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Bareket R, Schonberg MA, Schonmann Y. *BMJ Evidence-Based Medicine* 2020;25:12–14.

Abstract

The last decades saw remarkable change in the way healthcare professionals generate and consume medical knowledge. Information management technologies have evolved considerably, yet medical publications continue to use a referencing system that has changed very little since the turn of the 20th century. Research suggests that up to one in five referenced claims quotes the original text inaccurately. Many authors, perhaps inadvertently, contribute to this process by citing non-primary data and amplifying the errors of their predecessors. Erroneous claims are propagated, accumulate into false belief systems and generate inaccurate knowledge. Updating the referencing system to provide additional information to support each referenced claim (eg, the location of the referenced statement in the original text and the nature of that text) could, perhaps, address this cycle of inaccuracy. We believe such changes in the referencing system would prompt authors to rigorously verify referenced claims and provide readers with context to inform a critical evaluation of the text. We detail our proposal for changes in the notations used for referencing, as well as in the information provided within reference lists. We also discuss some barriers and solutions to the adoption of our proposal.

Introduction

John Ioannidis wrote in a frequently quoted 2005 paper that 'It can be proven that most claimed research findings are false'.¹(pg-696) He reasoned this was due to several mechanisms, including non-replication of research discoveries, bias in study design, analysis or presentation, low prestudy odds and conflicts of interest. In addition, he discussed the effect of biased prevailing assumptions. Ioannidis suggested that, when assessing the validity of a published work, one must consider the prestudy probability that the proposed findings are true.¹(pg-701) This could be achieved through a critical reading of the background and discussion sections of a research paper and by reviewing the cited literature. Unfortunately, readers of medical literature will experience some frustration when trying to do so because often the full text of a reference supporting an interesting or novel claim does not include the promised information.

Despite overwhelming changes and progress made in the past decades in information management, medical publications still use a referencing

system formalised more than a century ago in the *The Chicago Manual of Style*,² currently in its 17th iteration.³ In a changing world, with exponentially increasing amounts of data, this system might be obsolete and perhaps even an obstacle to medical progress.

Quotations, references and citations

The terms 'quotation', 'reference' and 'citation' are often used interchangeably but carry distinct meanings. Authors of academic texts often draw on the work of previous researchers and writers. They are, therefore, expected to clarify the sources of claims or statements included in their paper by attributing them to the earlier original work (ie, a reference). The earlier paper, which contained the claim in question, is then said to have a citation from the new paper. A quotation, strictly speaking, is a direct reproduction of someone else's text (or occasionally one's own text from elsewhere), included in inverted commas. Throughout this paper, however, we will include any repetition or paraphrasing of earlier ideas and claims in the term 'quotation'. We will reserve the term 'reference' to the more technical aspects of the citation process (ie, to describe the notations used to signify an attributed claim and the bibliographic details that are provided).

A Cochrane meta-analysis examined studies addressing quotation and referencing accuracy. A major quotation error was generally defined as a seriously misleading change to the original quoted claim and a major referencing error (referred to as a 'citation error' in that review) was generally defined as one that prevented or seriously obstructed the identification or retrieval of the reference. The authors concluded that approximately 20% of references in the medical literature were inaccurate quotes.⁴(R-7) These findings were replicated in a later meta-analysis by Jergas and Baethge,⁵(D-2) who further noted that 'there is an error in the proper sense in every fifth quotation, roughly half of them so severe that they are not at all in accordance with what the authors claimed: in an average article with 50 references, based on our figures, about six are completely wrong'.

Misleading quotations can be repeatedly cited to form a false belief-system, resulting in false-knowledge (ie, citation bias). Greenberg showed this in the field of neurology,⁶(D-2) and the same phenomenon was described by Harzing in a non-medical scientific field.⁷(F-1) Examples were also published in medical fields such as gastroenterology, anaesthesia, radiology, ophthalmology,

orthopaedics, public health, surgery, gynaecology, dermatology, emergency medicine, paediatrics, psychiatry, nephrology and infectious diseases.^{4(AmI-2,1)5(T-1)} We believe that improved, more informative references could result in fewer quotation errors and reduce citation bias.

The generation of false knowledge

Greenberg analysed papers that addressed the belief that a protein, beta-amyloid, plays an aetiological role in the injury caused to skeletal muscles in inclusion body myositis. He identified 242 papers that addressed that claim, of which 12 were primary research papers (ie, papers reporting on relevant original research findings), and the rest were reviews, cell culture or animal models, or other non-experimental reports. While half of the primary research papers refuted or weakened the hypothesis that beta-amyloid had an aetiological role, the supportive papers went on to receive 94% of the primary data citations. No papers refuted or critiqued the six original research papers that did not support the role of beta-amyloid in inclusion body myositis; they were simply ignored. Analysis of the network of citations that connected the 242 papers on this subject showed that 95% of all citation paths passed through four narrative review papers supporting the claim, all written by the same authors.^{6(R-10)} This created a lens effect, amplifying a supportive claim, and making other claims disappear, rather than discussed.

Greenberg suggested a framework for understanding how a false belief system is propagated. He identified several mechanisms by which improper quotations fail to represent the existing evidence and create false-knowledge, which is later amplified through further quoting.^{8(T-1)} First, authors might systematically neglect to cite previous papers that contrast with their claim ('citation bias'). Second, quotations could be flawed in several ways: misrepresenting the idea or claim expressed in the original paper ('citation diversion'); citing a hypothesis as a fact ('citation transmutation'); representation of abstracts as peer-reviewed papers ('back-door knowledge'); citing papers that do not actually address the referenced statement ('dead-end citations'); and statements in titles of papers that are not supported in the actual content ('title invention'). Third, and perhaps most influential, narrative reviews, editorials and other opinion pieces are sometimes mistaken for primary sources of evidence. Similarly, statements in the background and discussion sections of research articles are sometimes regarded as primary knowledge. Authors citing papers that do not contain primary data amplify and expand existing false belief systems, effectively erasing existing knowledge. As

more and more information accumulates, reviews become essential in organising existing data. However, only if systematically performed, using such frameworks as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement,⁹ can they be regarded as primary sources of knowledge.

Treating patients, allocating research resources and subjecting study participants to interventions on shaky grounds is just not good enough. Accurate quotations are essential for the scientific process, as a tool to assess the validity and prestudy probability of a research question, in providing a meaningful context for interpretation and ultimately as means to avoid the accumulation of false-knowledge.

Suggested solution

Several authors previously brought forward suggestions for addressing inaccurate quotations, but they mostly focused on guidance for authors and reviewers and not on actual systematic changes.^{5(D-20)} While relying on the scientific integrity of the authors and editors is a cornerstone of medical research and academic writing, we feel that additional measures could add context and facilitate critical reading. We propose a change in three aspects of referencing: clarifying the nature of the cited paper, noting the place of the quoted claim in the original article and explicitly stating whether the claim refers to primary research results (see examples in table 1).

1. Nature of the cited paper: the reference list would be formatted to include basic information on the nature of the cited paper and not just bibliographic information. This information should clarify the study design (ie, randomised controlled trial, cohort, case-control, narrative review, systematic review, meta-analysis, editorial and so on) if not already in the title, the sample size, the year when data collection was completed and the follow-up period.
2. Place of the quoted claim: in the text of the article, following the reference number, authors should point the reader to the relevant specific location in the paper they cited. The notation could use the ubiquitous introduction, methods, results and discussion structure, and add a paragraph number. For example, 'R-2' asserts that the statement quotes an idea presented in the second paragraph of the results section (see table 2). We do not mean to imply that certain locations are more reliable than others but merely to ease the process of locating and assessing the original claim.
3. Source of the data: in the text of the article, following the reference number, the authors should specify whether they

Table 1 Examples of claims and how they should be referenced

Claim	Reference
Smoking five cigarettes a day, or less, is associated with over a half of the cardiovascular risk associated with smoking a pack a day. This effect is seen in both men and women, and across different age groups. ^{1*(Fig2)}	1. Hackshaw, Allan, Joan K. Morris, Sadie Boniface, Jin-Ling Tang, and Dušan Milenković. 'Low Cigarette Consumption and Risk of Coronary Heart Disease and Stroke: Meta-Analysis of 141 Cohort Studies in 55 Study Reports'. <i>BMJ</i> 360 (January 24, 2018): j5855. https://doi.org/10.1136/bmj.j5855 . (n=3 070 000, Follow-up: 10–21 years)
Patients with anti-NMDAR encephalitis have an increased risk for developing neuroleptic malignant syndrome as a side effect of neuroleptic treatment. ^{2(Pg-848)}	2. Dalmau, Josep, and Francesc Graus. 'Antibody-Mediated Encephalitis'. <i>New England Journal of Medicine</i> 378, no. 9 (March 1, 2018): 840–51. https://doi.org/10.1056/NEJMra1708712 . (Narrative review)
Zika virus PCR assays were positive in autopsies of babies with brain abnormalities whose mothers had fever and rash during their pregnancy. ^{3*(R1-2)}	3. Martines, Roosecelis Brasil, Julu Bhatnagar, Ana Maria de Oliveira Ramos, Helaine Pompeia Freire Davi, Silvia D'Andretta Iglezias, Cristina Takami Kanamura, M. Kelly Keating, et al. 'Pathology of Congenital Zika Syndrome in Brazil: A Case Series'. <i>The Lancet</i> 388, no. 10047 (August 27, 2016): 898–904. https://doi.org/10.1016/S0140-6736(16)30883-2 . (Case series, n=5, 2015–2016)

Examples.

Table 2 Abbreviations for the location of the quoted claim

Location	Abbreviation
Introduction	I
Methods	M
Results	R
Discussion	D
Table	T
Figure	F
Supplement	S
Page	Pg

cite primary data, (ie, research results). The notation could be an addition of an asterisk after the reference number to signify primary data.

Admittedly, ours is not the only possible way to address the issue of misleading quotations. For example, journals could encourage writers to include detailed descriptions of the sources they cite (eg, 'A recent randomised controlled study concluded that...'). Alternatively, journals can adopt the use of an annotated bibliography (ie, including brief textual descriptions of the cited source). Such solutions, however, might interfere with the flow and readability of the text and lack the structure and coherence of a more systematic approach. The role of journal editors and reviewers should also be mentioned in this context, as it is their formal responsibility to guarantee the accuracy of referenced claims. Perhaps reviewers should be encouraged to examine reference lists and 1–2 original sources when reviewing manuscripts. While the adoption of our proposal would not protect against frank misconduct, the changes we suggest in the referencing system could facilitate the work of editors and reviewers, as well as increase the transparency and validity of scientific communication.

A major disadvantage for authors and researchers would be the added burden of composing detailed reference entries. However, the process could easily be automated. Most authors use citation management software to generate formatted reference lists; the details and descriptors of each reference could be incorporated into the existing meta-data within these programmes. This drawback could be outweighed by the potential benefits: authors would be prompted to reflect on the accuracy of claims, and readers could critically assess what they read. The practice of citing low-grade evidence would be discouraged, and most of the mechanisms suggested by Greenberg for inaccurate knowledge would be addressed. Finally, references to non-primary sources of data would be highlighted. This approach would not solve the issue of citation bias (or confirmatory bias), where an author preferentially selects the quotes that support their claim but could reduce the tendency of such claims to be amplified by future repeated citing. On the long run, widespread use of these proposed changes could result in general improvement of quality and reliability of the generated knowledge.

Contributors RB is a family physician; he is head of 'Plotkin' clinic in Clalit Health Services (CHS; Israel's largest non-for-profit health provider), the secretary of the Israeli Society for Reduction of Overdiagnosis (ISROD) and a researcher at the Office of the Chief Physician at CHS and at the Sackler department of Medicine in the Tel Aviv University. MAS is an associate professor of Medicine at Beth Israel Deaconess Medical Center and at Harvard Medical School. YS is a family physician at CHS, a researcher at the Office of the Chief Physician at CHS and the Department of Noncommunicable Disease Epidemiology at the London School of Hygiene & Tropical Medicine. The article arose from discussion and observations of RB and YS regarding the difficulties and challenges of interpretation and application of current medical evidence. RB conceived the original concept. Literature search and drafting of the manuscript were performed by RB and YS. MAS contributed to the development of the idea. All authors participated in critically revising of the manuscript, approved it, and agree to be accountable for all its aspects. RB and YS are the guarantors of this article.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs

Ronen Bareket <http://orcid.org/0000-0003-1349-5885>

Yochai Schonmann <http://orcid.org/0000-0002-4483-2277>

References

- Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124–701.
- University of Chicago. *Manual of style : being a compilation of the typographical rules in force at the University of Chicago Press, to which are appended specimens of types in use*: University of Chicago Press, 1906.
- Webster G. The Chicago Manual of Style, Seventeenth Edition. *Publishing Research Quarterly* 2017;33:495–6.
- Wager E, Middleton P. Technical editing of research reports in biomedical journals. *Cochrane Database Syst Rev* 2008;MR000002.
- Jergas H, Baethge C. Quotation accuracy in medical journal articles—a systematic review and meta-analysis. *PeerJ* 2015;3:e1364.
- Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. *BMJ* 2009;339:b2680.
- Harzing A-WK. The persistent myth of high expatriate failure rates. *The International Journal of Human Resource Management* 1995;6:457–74.
- Greenberg SA. Understanding belief using citation networks. *J Eval Clin Pract* 2011;17:389–93.
- Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.