Reply

ISSN 0946-2104

DOI 10.5414/TEX01495
e-pub: May 17, 2017

REPLY TO LETTER

RE: Seifert M. How accurate are references in *Trace Elements and Electrolytes*? Trace Elem Electrolytes. 2017; *34:* 137-138.

Robin Dunford[1,2], Sylvia Izzo Hunter[2,3], and Bruce Rosenblum[2]

[1]Dunford Consulting, Wantage, UK, [2]Inera Incorporated, Belmont, MA, USA, and [3]All in the Words, Toronto, ON, Canada

Sir, – We note with interest the results of Dr. Seifert's study of reference accuracy in *Trace Elements and Electrolytes (TEE)*, and his finding that the error rate is low relative to previous studies of reference accuracy in medical journals [1]. Publishers increasingly use technology-based solutions to aid checking and correcting references [2]. The Editorial Office of *TEE* uses eXtyles software from our company, Inera, to automatically verify and correct references against PubMed and Crossref data.

Such software may be one factor among many that improve reference accuracy. Other factors may include a higher level of diligence among this community of authors and reviewers, and the editorial work of *TEE*'s professional staff.

Although we are pleased that the error rate reported by Dr. Seifert is low, our research and development team has carefully examined the 31 references identified as containing major errors, in an effort to understand the source of those errors and whether future software improvements could further reduce the error rate. The errors mostly fall into three groups:

1. References that are not indexed in PubMed or Crossref;
2. Author-supplied reference entries so different from the online source that our software could not match them to an online record; and
3. Instances in which the reference entry matches the online source(s), but the online source differs from the "version of record".

References in the first two error groups are more time consuming for an editorial office to check, and some publishers may be satisfied if the overwhelming majority of references link to at least one of PubMed or Crossref. Having met this standard, which allows readers to locate the source, a publisher may make the business decision that further improvements in accuracy are not cost effective or mission critical.

For references *not* linked to these online resources, citation accuracy is arguably even more important, because readers must conduct their own searches to locate these materials. Errors in reference entries may make such searches more difficult, or even impossible. Indeed, the largest group of errors in Dr. Seifert's dataset occur in references to sources not indexed in either PubMed or Crossref. This brings us to another key point: no software, no matter how sophisticated, can entirely replace careful human attention to accuracy in citations.

Errors in group 3 above require more extensive examination. Software validation uses different data sources from those available to a human researcher. The "version of record", correctly used by Dr. Seifert, is the printed journal or online PDF, but it is not practical or even possible for the software to check against this definitive version, since no standard method exists for automated queries; instead, the software uses machine-readable PubMed and Crossref metadata as a proxy. In theory, these sources contain the same information as the version of record, as the metadata is supplied by the publisher or is curated by librarians from the published article. In practice, metadata matches version of record in the majority of cases, but a small number of records contain discrepancies.

Our analysis found that most Crossref- and PubMed-linked references with major errors had metadata errors at both services, which suggests the errors likely originated at the respective publishers. The most systematic error occurred in author name segmentation (the separation of name elements into given name(s) and surname(s)). Publishers

and their vendors may need to review and improve their processes for capturing correct author name information in the publication workflow and verifying that all metadata is correct, including not just the spelling but the order and segmentation of author names.

The availability of online metadata, and of software allowing editorial offices to use this metadata more cost effectively, presents new opportunities for publishers to check and correct references without time-intensive labor. Further study is necessary to determine the extent to which such software has contributed to the results reported by Dr. Seifert. Future research could isolate different factors and extend Dr. Seifert's data set to include a control, such as a volume of *TEE* published before the software was introduced. Independent of such further studies, as software developers our goal is to build on Dr. Seifert's encouraging findings and provide software for publishers to achieve even lower reference error rates.

Review of the errors reported by Dr. Seifert shows that work remains to be done to further improve reference accuracy. The advent of software tools to correct references does not absolve authors from responsibility for reference quality, nor editors from the need to check and correct references that are not easily linkable to online sources – the most challenging ones for readers to find. Finally, we encourage publishers to institute or extend processes to carefully check the quality of metadata they provide to online indexers and metadata providers. Improving reference accuracy, including metadata uploads, benefits all participants in the scholarly communication ecosystem.

References

[1] *Seifert M*. How accurate are references in Trace Elements and Electrolytes? Trace Elem Electrolytes. 2017; *34:* 137-138. CrossRef

[2] *Meyer CA*. Reference accuracy: best practices for making the links. J Electron Publ. 2008; *11*: https://doi.org/10.3998/3336451.0011.206. CrossRef

Correspondence to:
Bruce Rosenblum
Inera Incorporated
19 Flett Rd.
Belmont, MA 02478, USA
bruce@inera.com