

LETTER

Open Access



Response to ‘Increasing value and reducing waste in data extraction for systematic reviews: tracking data in data extraction forms’

Jens Jap^{1*}, Ian J. Saldanha², Bryant T. Smith¹, Joseph Lau¹ and Tianjing Li²

Abstract

This is a response to a Letter. Data abstraction is a time-consuming and error-prone systematic review task. Shokraneh and Adams categorize available techniques for tracking data during data abstraction into three methods: simple annotation, descriptive addressing, and Cartesian coordinate system. While we agree with the categorization of the techniques, we disagree with the authors’ statement that descriptive addressing is a PDF-independent method, i.e., any sort of descriptive addressing must reference a specific version of PDF file and not just any PDF of said report. Different versions of PDFs of the same report might place text and tables on different locations of the same page and/or on different pages. Consequently, it is our opinion that any kind of source location information should be accompanied by the source or linked by an intermediary service such as the Data Abstraction Assistant (DAA).

We read with great interest Shokraneh and Adams’ letter pertaining to data abstraction (or “data extraction”) during systematic reviews [1]. We agree with the authors that data abstraction is perhaps the most time-consuming task during systematic reviews and one that is error-prone [2–5]. Manual data abstraction, which is largely the current norm in the systematic review enterprise, is likely not sustainable in the long run. Software tools that help tracking of data to published reports of studies (i.e., PDFs) have the potential to greatly reduce the time spent and errors inherent to the data abstraction process [6]. We developed Data Abstraction Assistant (DAA), a software tool with this potential. By recording the exact location and mapping this location to the data entered into extraction forms, DAA could reduce errors and time spent reviewing extracted data [6].

Shokraneh and Adams categorize available techniques for tracking data into three methods: simple annotation, descriptive addressing, and Cartesian coordinate system

[1]. The authors describe the second method, i.e., descriptive addressing, as one where the data abstractor abstracts data and notes each data point’s source of information (“address”) in the PDF, using the page, paragraph, line, table, figure, box, and/or headline numbers.

While we agree with the categorization of the techniques, we disagree with the authors’ statement that descriptive addressing is “PDF-independent.” In our experience, descriptive addressing indeed is dependent on the version of the PDF used; different versions of PDFs of the same report might place text and tables on different locations of the same page and/or on different pages. Consequently, it is our opinion that any kind of source location information should be accompanied by the source or linked by an intermediary service such as DAA. We developed DAA to facilitate data tracking between data abstraction forms and PDFs, thereby possibly reducing errors and saving time [6]. DAA allows users to mark and record the exact location of information found on a PDF. The locations are linked to data elements on a data extraction form. DAA enables users to create a link between information extracted and its source location. We are currently analyzing the results of a randomized controlled trial that formally evaluates the

* Correspondence: jens_jap@brown.edu

¹Center of Evidence Synthesis in Health, Brown University School of Public Health, Providence, USA

Full list of author information is available at the end of the article

effectiveness of DAA (compared with standard data abstraction approaches) in improving these outcomes (i.e., error rates and time).

The use of DAA would not solve the challenge that copyright poses in sharing PDFs. However, by serving as an intermediary, linking the abstracted data and the exact location in the PDF source, DAA facilitates the efficient tracking of abstracted data.

Acknowledgements

Not applicable.

Funding

This work is funded by a PCORI contract (Contract Number: ME-1310-07009).

Availability of data and materials

Not applicable.

Authors' contributions

JJ conceived this manuscript. JJ and IJS drafted this manuscript, and all authors provided input and approved the final version of this manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors provide consent for publication of this manuscript.

Competing interests

All authors are investigators of a Patient-Centered Outcomes Research Institute (PCORI)-funded project (Contract Number: ME-1310-07009) that consists of developing DAA and conducting a randomized controlled trial to evaluate its effectiveness.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Center of Evidence Synthesis in Health, Brown University School of Public Health, Providence, USA. ²Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA.

Received: 5 September 2017 Accepted: 10 January 2018

Published online: 25 January 2018

References

1. Shokraneh F, Adams CE. Increasing value and reducing waste in data extraction for systematic reviews: tracking data in data extraction forms. *Syst Rev*. 2017;6:133.
2. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol*. 2006;59(7):697–703.
3. Gøtzsche PC, Hróbjartsson A, Maric K, Tendal B. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA*. 2007;298(4):430–7.
4. Jones AP, Remington T, Williamson PR, Ashby D, Smyth RL. High prevalence but low impact of data extraction and reporting errors were found in Cochrane systematic reviews. *J Clin Epidemiol*. 2005;58(7):741–2.
5. Horton J, Vandermeer B, Hartling L, Tjosvold L, Klassen TP, Buscemi N. Systematic review data extraction: cross-sectional study showed that experience did not increase accuracy. *J Clin Epidemiol*. 2010;63(3):289–98.
6. Saldanha IJ, Schmid CH, Lau J, Dickersin K, Berlin JA, Jap J, Smith BT, Carini S, Chan W, De Bruijn B, Wallace BC, Hutfless SM, Sim I, Murad MH, Walsh SA, Whamond EJ, Li T. Evaluating data abstraction assistant, a novel software application for data abstraction during systematic reviews: protocol for a randomized controlled trial. *Syst Rev*. 2016;5(1):196.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

