Towards automatic data extraction from clinical research reports: a case study of a systematic review of oral pain relief

Linh Hoang[1], Tanja Bekhuis[2], Jodi Schneider[1]

[1] School of Information Sciences, University of Illinois at Urbana-Champaign [2] Department of Biomedical Informatics, University of Pittsburgh

School of Information Sciences
The iSchool at Illinois

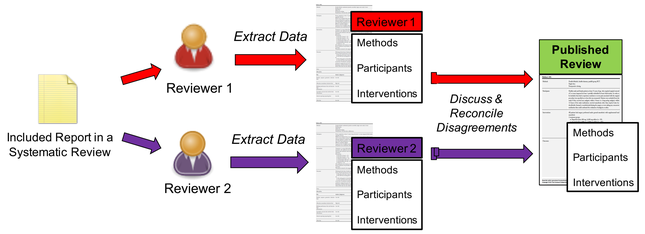
1. Introduction

Systematic Review:

- Systematic review is a type of literature review. In healthcare, to answer a particular clinical question, all available evidence is synthesized into a single systematic review.
- The systematic review process includes a series of steps. Data extraction from clinical research reports is one of the most time-consuming steps.

Motivation: In health care, it takes a long time for new treatments to move from clinical studies into practice: perhaps an average of 17 years [Balas et al., 2000].

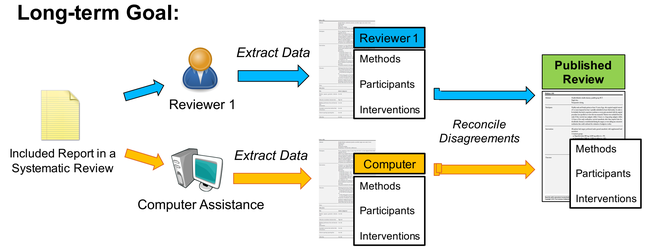
Data Extraction in Systematic Review:



***CURRENT BEST PRACTICE**: At least 2 reviewers independently extract data from the included research reports. They reconcile differences to reach consensus before synthesizing the evidence.*

Problem: The data extraction step is almost always performed manually. Data extraction is very time-consuming [Tsafnat et al., 2014] yet methodological errors may cause problems with the review's conclusions [Lundh et al., 2009].

Long-term Goal:



***PROPOSED SEMI-AUTOMATION:** A semi-automated system could support a single reviewer during data extraction. Differences in information extracted by a human reviewer and a computerized system could be displayed. The reviewer decides on the consensus version.*

Goal: Our long-term goal is to help reviewers synthesize the literature quickly and accurately by developing a semi-automatic support system for data extraction.

2. Related Work

RobotReviewer *[Wallace et al., 2016]*

- RobotReviewer is an automatic data extraction system. It uses machine learning and natural language processing to extract data from clinical research reports.
- RobotReviewer extracts 3 key elements (Participant, Intervention & Outcome) from the full-texts of the clinical research reports.

3. Research Questions and Methods

Questions:

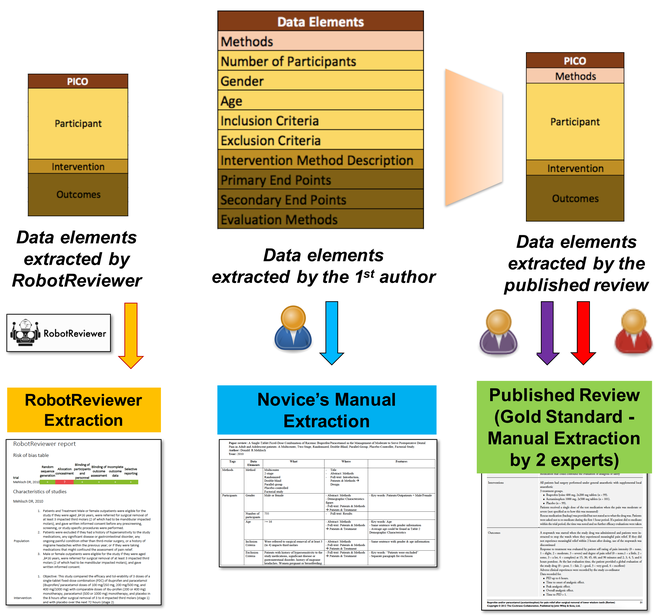
- How does RobotReviewer's data extraction compare to systematic reviewers' data extraction?
- How does RobotReviewer's data extraction compare to a single novice reviewer's data extraction?

Methods:

An in-depth case study of a single systematic review, a Cochrane Review about oral pain relief [Bailey et al., 2013], which synthesizes 6 clinical research reports.

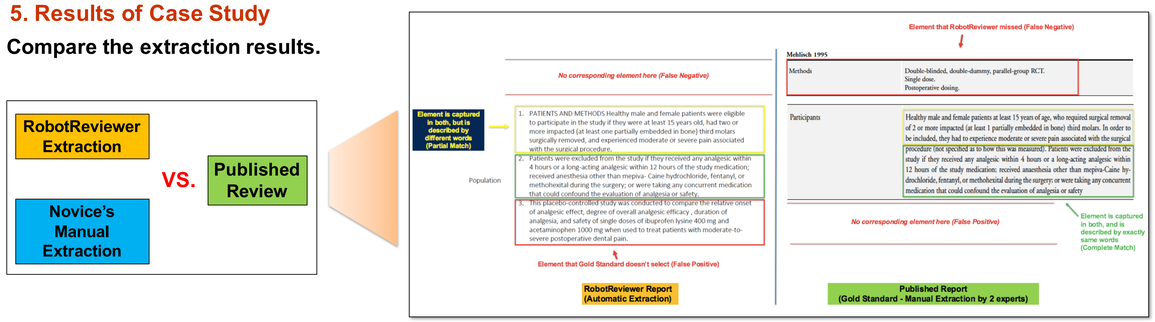
- Manually extract data elements from the 6 included reports.
- Run RobotReviewer on the 6 included reports.
- Compare the novice's manual extraction and the RobotReviewer's extraction with the published review as a gold standard.

4. Data Extraction



5. Results of Case Study

Compare the extraction results.



Comparing data elements with the published review: complete agreement, partial agreement, or no agreement.

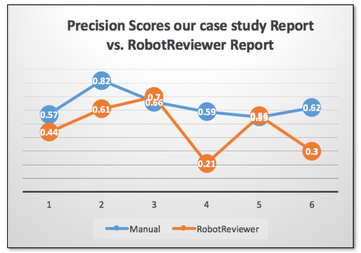
Calculate information retrieval metrics.

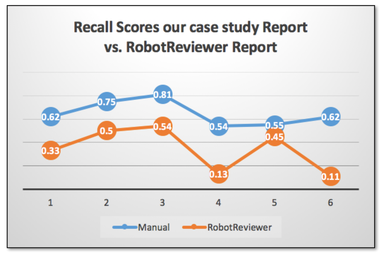
Precision: The percentage of data elements that are correctly identified in RobotReviewer's extraction or the novice's manual extraction.

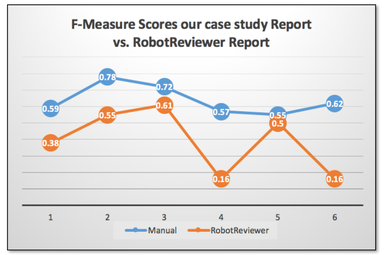
Recall: The percentage of data elements that are correctly identified, comparing with the published review.

F-Measure: A weighted average of Precision and Recall.

Paper	Annotation	# False Positives	#False Negatives	#Complete Matches	#Partial Matches	Precision	Recall	F-Measure
Paper 1	Manual (v2)	3	2	5	6	0.57	0.62	0.59
	RobotReviewer	4	7	3	2	0.44	0.33	0.38
Paper 2	Manual	1	2	8	2	0.82	0.75	0.78
	RobotReviewer	3	5	5	1	0.61	0.50	0.55
Paper 3	Manual	4	1	9	3	0.66	0.81	0.72
	RobotReviewer	1	4	5	4	0.70	0.54	0.61
Paper 4	Manual	3	4	5	3	0.59	0.54	0.57
	RobotReviewer	5	10	1	1	0.21	0.13	0.16
Paper 5	Manual	4	4	5	2	0.55	0.55	0.55
	RobotReviewer	3	5	4	2	0.56	0.45	0.50
Paper 6	Manual	4	4	7	2	0.62	0.62	0.62
	RobotReviewer	3	12	1	1	0.30	0.11	0.16
Average	Manual					0.63	0.65	0.64
	Robot Reviewer					0.47	0.34	0.39







6. Discussion

Data Extraction Results:

- Identify potential features for automating data extraction.
- Develop hypotheses about which features could be used to automate data extraction.

Case Study Results:

- RobotReviewer's extraction results are inconsistent.
- Performance measures for the novice's manual extraction are not as high as expected. However, they are consistent for the 6 studies included in the systematic review (P,R,F are all ranged from 0.55 – 0.8).

7. Limitations

- Small sample size.
- 3/6 articles from the same author, 4/6 articles were published in the same journal.
- The novice's manual extraction was not completely independent: data extraction from the first paper was done after looking at the published review.
- The evaluation was based on a list of data elements which were created by the 1st author.

8. Next Steps

- Increase the sample size by examining more systematic reviews and the clinical research reports they include.
- Consider other metrics for assessing data extraction quality.
- Draw on the PICO ontology [The Cochrane Collaboration, 2014-] to update the list of data elements to be extracted.
- Conduct an error analysis and study RobotReviewer's code to understand what works and where it goes wrong.
- Interview systematic reviewers to understand their expectations for automation.

References

- *Bailey E, Worthington HV, Van Wijk A, Yates JM, Coulthard P, Afzal Z. Ibuprofen and/or paracetamol (acetaminophen) for pain relief after surgical removal of lower wisdom teeth. The Cochrane Library 2013.*
- *Balas EA, Boren SA. Managing clinical knowledge for health care improvement. In: Bemmel J, McCray AT, editors. Yearbook of Medical Informatics 2000: Patient-Centered Systems. Stuttgart, Germany: Schattauer Verlagsgesellschaft mbH; 2000. p. 65-70.*
- *Lundh A, Knijnenburg SL, Jørgensen AW, Van dalen EC, Kremer LC. Quality of systematic reviews in pediatric oncology--a systematic review. Cancer Treat Rev. 2009;35(8):645-52.*
- *Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. Syst Rev. 2014;3:74.*
- *Wallace BC, Kuiper J, Sharma A, Zhu, MB, & Marshall IJ. Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision. Journal of Machine Learning Research 17 (2016) 1-25.*
- *The Cochrane Collaboration. PICO ontology [Internet]. London: The Cochrane Collaboration; c2014- [cited 2016 Oct 21]. Available from: http://data.cochrane.org/ontologies/pico/.*

Funded by training grant T15LM007059 from the National Library of Medicine/National Institute of Dental and Craniofacial Research and by startup funds from the University of Illinois School of Information Sciences.

ILLINOIS