

# Comparing articles identified as Randomized Controlled Trials: MEDLINE, Cochrane, and the RCT Tagger

Yogeshwar Kansara, Linh Hoang, Jodi Schneider

School of Information Sciences, University of Illinois at Urbana-Champaign

## Motivation

- Randomized Controlled Trials (RCTs) are considered the gold standard for knowledge about the efforts of medical treatments. In evidence-based clinical practice, identifying the Randomized Controlled Trials in search results can be useful for supporting clinicians in finding high-quality information.
- Our goal is to understand the impact of different approaches of identifying RCTs on information retrieval for systematic reviews by comparing articles identified as RCTs from three approaches: MEDLINE, Cochrane, and the RCT Tagger.

## MEDLINE

- A popular database of biomedical literature, produced by the US National Library of Medicine since the 1960's, often accessed through web-based interfaces such as Ovid or PubMed [1].
- Uses Publication Type "Randomized Controlled Trial" to index RCT articles [2].

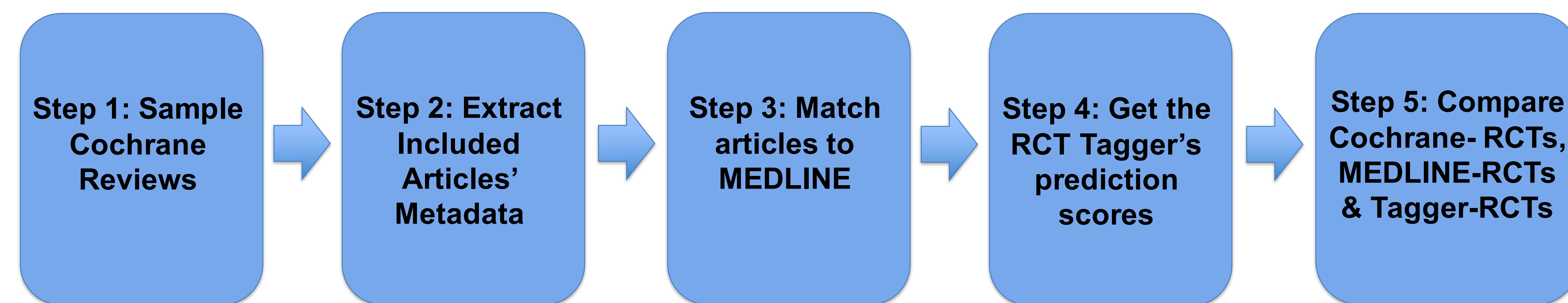
## Cochrane

- A leading provider of systematic reviews of medical research.
- Maintains their own library of trials, the Cochrane Central Register of Controlled Trials, which is designed to aid in information retrieval for systematic reviews [3].

## The RCT Tagger

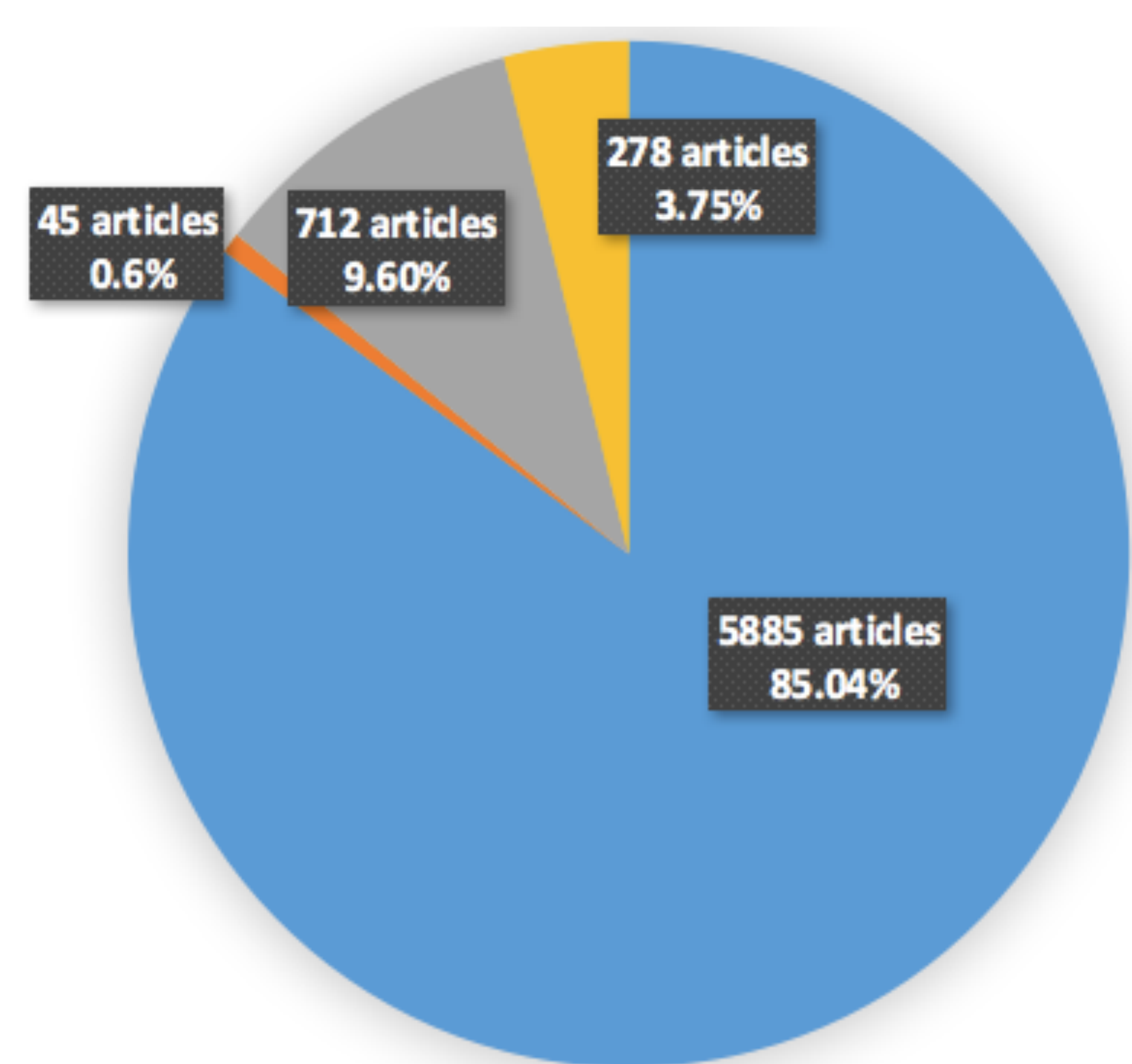
- A classifier, designed to identify RCTs.
- Uses text mining and machine learning to analyze the text in papers' titles and abstracts [4].
- Gives a prediction score from 0 to 1. For systematic reviewing, the recommendation is to screen articles with scores above 0.01 as possible RCTs [4].

## Case Study Design



## Results Summary

- **Step 1:** 895 reviews from the original dataset were randomly selected. 570 of these reviews were categorized as only including RCTs after human annotation of the reviews' inclusion criteria.
- **Step 2:** 10,192 unique included articles and their metadata were extracted from the reviews.
- **Step 3:** 7413 unique articles with PubMed IDs and Publication Types.
- **Step 4:** 6920 articles were processed by the RCT Tagger and got prediction scores.
- **Step 5:** We compared the three approaches for identifying whether or not these 6920 articles were RCTs:



Number (percentage) of articles	Cochrane-RCT	MEDLINE-RCT	Tagger-RCT (scores >.01)	Description of the category
5885 (85.04%)	Yes	Yes	Yes	All approaches agreed
45 (0.6%)	Yes	Yes	No	The RCT Tagger disagreed with Cochrane & MEDLINE
712 (9.6%)	Yes	No	Yes	MEDLINE disagreed with Cochrane & the RCT Tagger
278 (3.75%)	Yes	No	No	Cochrane disagreed with MEDLINE & the RCT Tagger

## Distribution of Tagger Scores

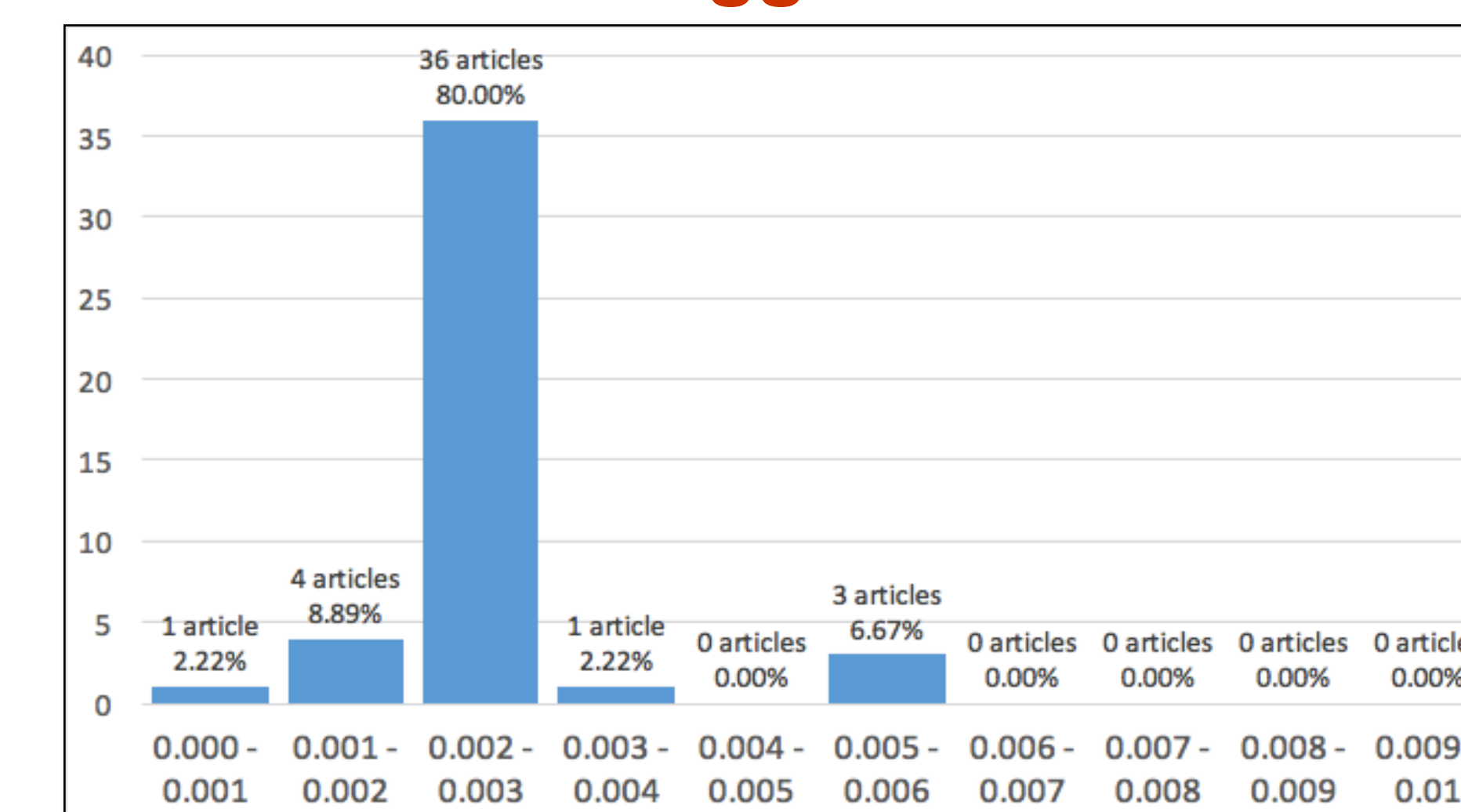


Fig. 1: Distribution of Tagger scores of "All but Tagger" articles (45 articles). Percentages relative to the category.

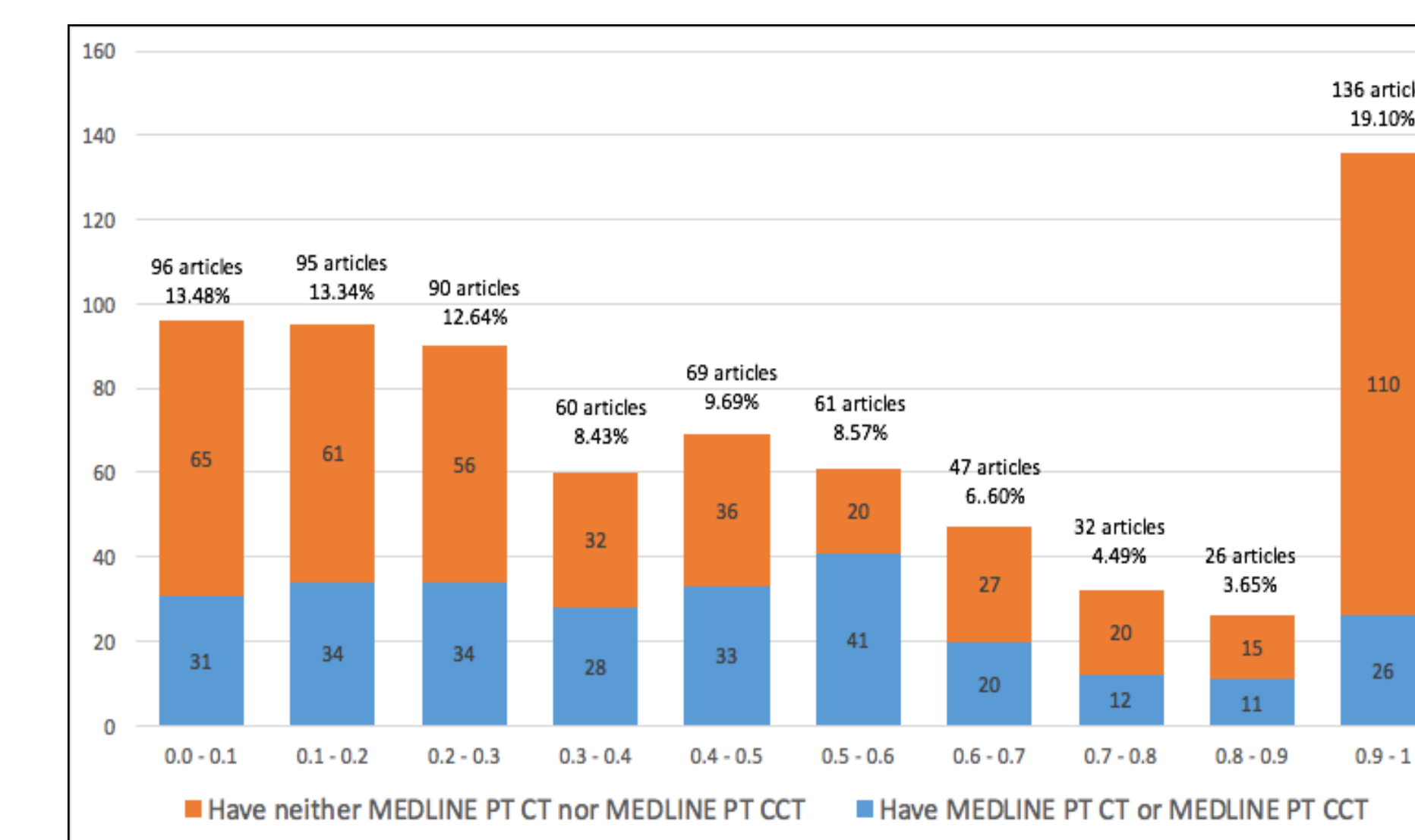


Fig. 2: Distribution of Tagger scores of "All but MEDLINE" articles (712 articles). Percentages relative to the category.

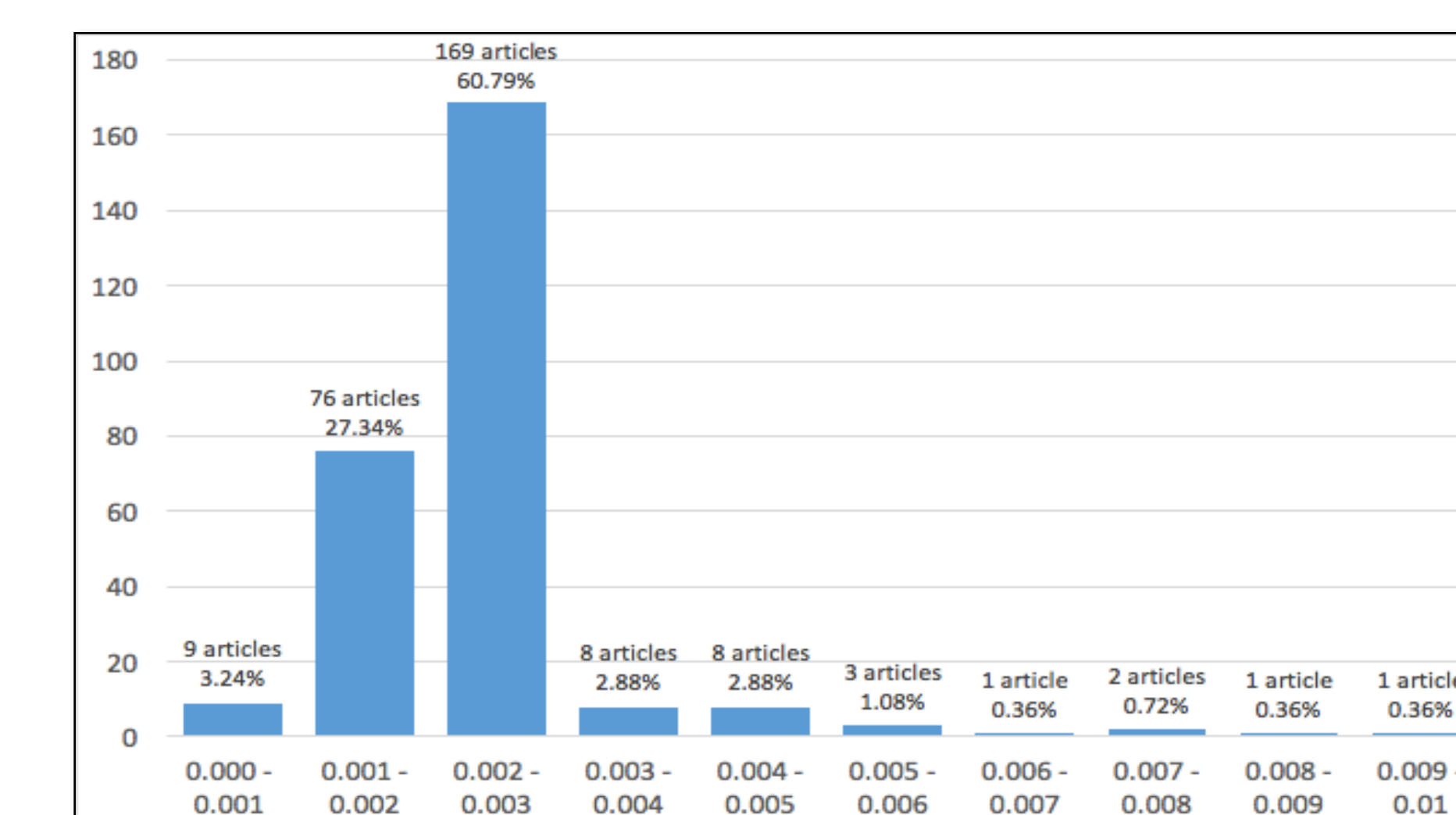


Fig. 3: Distribution of Tagger scores of "Cochrane Only" articles (278 articles). Percentages relative to the category.

## Discussion

### Comparing Cochrane, MEDLINE, and RCT Tagger's RCTs.

- The three different approaches are all somewhat differently defined, and all three are prone to errors of their own.
  - Previous research on Tagger has estimated that MEDLINE has about a 7% error rate in Tagging RCTs4 [4].
- None of these is a gold standard relative to the others.
  - The inter-rater agreement from previous research on Tagger was reported as 72% [4].

### The RCT Tagger Evaluation:

- Tagger's error is only 45/5930 which is less than 1%. Of 5930 papers considered both Cochrane-RCTs and MEDLINE-RCTs, only 45 were not considered Tagger-RCTs.
- Abstracts may not have enough information to classify articles. "Random allocation" information was only in the whole paper, not the abstracts, for 20 of the 45 articles where RCT Tagger prediction conflicted with both Cochrane and MEDLINE.
- Missing abstracts caused some data to not be processed. The RCT Tagger did not process 493 articles included in Cochrane reviews (6.6% of our sample) either because they had no abstract in PubMed (491 articles, 461 English, 30 non-English) or because the full-text was not in English (2 articles had English abstracts in PubMed even though their full-text was in Chinese).

## Future Work

- Suggest improvements for the RCT Tagger based on an error analysis.
- Evaluate how RCT classification impact the systematic reviewing process.

## References

1. National Library of Medicine. MEDLINE Fact Sheet. Last updated July 23, 2018. <https://www.nlm.nih.gov/bsd/medline.html>
2. Glanville JM, Lefebvre C, Miles JN, & Camosso-Stefinovic J (2006). How to identify randomized controlled trials in MEDLINE: ten years on. J Med Libr Assoc., 94(2), 130.
3. The Cochrane Collaboration. The Cochrane library. Chichester, UK: John Wiley & Sons, 2005. <http://www.thecochranelibrary.com>
4. Cohen AM, Smalheiser NR, McDonagh MS, Yu C, Adams CE, Davis JM, Yu PS. Automated confidence ranked classification of randomized controlled trial articles: an aid to evidence-based medicine. J Am Med Inform Assoc. 2015 Feb 5;22(3):707-17.

## Acknowledgements

This work was funded by the NIH/NLM under R01LM010817

Thank you to undergraduates Xiaoru Dong and Jingyi Xie for assistance annotating inclusion criteria to find RCT-only Cochrane reviews.