

Original Article

The effectiveness of methodological elements in ranking results by relevance

Adela Asaadi Shally¹, Hajar Sotoudeh^{2*}, Javad Abbaspour³

¹ PhD, Knowledge and Information Sciences, Faculty of Education and Psychology, Shiraz University, Shiraz, Iran.

² Professor, Faculty of Education and Psychology, Shiraz University, Shiraz, Iran.

³ Assistant Professor, Faculty of Education and Psychology, Shiraz University, Shiraz, Iran.

ARTICLE INFO

Corresponding Author:
Hajar Sotoudeh
e-mail addresses:
sotoudeh@shirazu.ac.ir

Received: 2/Oct/2022
Modified: 19/Dec/2022
Accepted: 21/Dec/2022
Available online: 11/Jun/2023

Keywords:

Methodological elements
Cochrane system review articles
Relevance
Information retrieval
Methodological validity
Reviewers' opinions
Cochrane Database
Clinical Trial

ABSTRACT

Introduction: Methodological validity is one of the aspects of quality. Methodological elements are parts of the text of articles that deal with research methodology. The purpose of this study was to determine the contribution and role of methodological elements in explaining the relationship between evidence and questions.

Methods: This semi-experimental study employed a one-group pretest-posttest design. The research population consisted of clinical trial articles included in the meta-analysis of Cochrane systematic review articles. The sampling method employed was purposeful, whereby systematic review articles containing at least 50 related clinical trial articles retrieved by the retrieval system were selected as the research sample.

Results: The results of the paired t-test showed that the difference in the average nDCG score across all four groups was negative at all points of accuracy. The highest average difference (-0.064) was observed for the basic and standard methodological elements in the abstract at accuracy point 10 (the tenth document in the retrieved results), while the lowest average difference (-0.021) was observed for the basic methodological elements in the abstract at the 50th and 70th accuracy points.

Conclusion: The findings of this research showed that methodological elements, whether independently or to expand the abstract, do not affect the ranking of relevance results or may even have a negative effect. In other words, the occurrence of methodological elements in the text or their weighting can reduce relevant results.

Extended Abstract

Introduction

Retrieving medical information is crucial from various perspectives. Health professionals are motivated to search databases for purposes such as learning, teaching, and reviewing research backgrounds, writing systematic reviews and meta-analyses, and answering clinical questions. Classification of articles on the basis of different research methods, automatic extraction of certain parts of articles, extraction of article abstracts, and evaluation of article validity are integral components of ongoing projects aimed at reducing the time and cost associated with information retrieval for users. The quality of information is one of the factors influencing relevance judgments. It garners attention not only from general users of databases but also from a diverse range of experts, including judges, researchers, and professionals across various fields. The quality of information encompasses extremely complex dimensions. Methodological validity, in particular, is one of the aspects of quality that has considerable importance in scientific articles as it ensures closer alignment with reality and fosters trust in the reported results. Authoritative articles exhibit a higher level of caution in presenting their findings. [1-5] Previous research on text mining and natural language processing in the context of evidence-based medicine has shown that most studies in this field focus on the categorization of articles according to the PICO (patient, intervention, comparison or control, and outcome) structure, [6-11] automatic selection of relevant articles, [12-14] recognition of medical articles on the basis of their similarity, [15] extraction of important words from scientific article sections, [7,11,16-22] identification and classification of article types, [9,23] and summarization of article texts [10,24-27].

These studies have employed natural language processing techniques and text mining to address the above-mentioned challenges. However, upon a review of the literature, it becomes evident that there is a dearth of research about the role of methodological elements in determining article relevance and validity. Therefore, the present study aimed to measure how effectively the methodological elements of article texts could be used as a tool to determine paper relevance. By doing so, this research endeavored to bridge the knowledge gap in this field.

Methods

The present study employed a pretest–posttest quasi-experimental design to achieve the research objectives. The independent variable was textual similarity, while the dependent variable was expert relevance. The basic methodological elements consisted of text words and the qualitative reviewer’s opinion about the clinical trial articles included in the meta-analysis of the Cochrane systematic review articles. [28-32] A study by Asadi et al. [33] in 2020 showed that Cochrane reviewers’ opinions addressed both the topic of the article and the research method. Therefore, in this study, the words of Cochrane reviewers’ opinions (considered as basic methodology elements) were incorporated as the primary indicators of methodological validity. Furthermore, to enrich the methodological vocabulary within the Cochrane reviewers’ comments, we derived synonyms of the words found in their opinions from medical texts, primarily focusing on books related to medical research methods. In the current research, basic and standard methodological elements were identified and tested across four groups. The first group comprised common words found in both the Cochrane reviewers’ opinions and article abstracts.

Subsequently, the common words from the Cochrane reviewers' opinions, along with relevant medical text words, were extracted and added to the abstracts. The remaining two groups involved incorporating all the words from the Cochrane reviewers' opinions into the abstract text and adding the words from the Cochrane reviewers' opinions and those derived from medical texts to the abstract text. In research pertaining to information retrieval and efficacy evaluation of the retrieved result ranking, the presence of certain factors such as questions, obtained results, and their relevance to the questions is essential. Systematic review articles are written in response to a clinical question, and therefore, each systematic review article is considered a question. To write systematic review articles, researchers conduct extensive searches across several databases and select articles related to the topic at hand. Therefore, all the articles used in the meta-analysis section of the systematic review are deemed pertinent to the clinical question, thus serving as the population for the current study. In this study, in order to sample the research community and prepare the test set, we initially utilized a Python program to download the hypertext markup language (HTML) text of 3,793 open access systematic review articles from the Cochrane database (review articles from the beginning to the end of 2018). These articles included a compilation of clinical trial articles used in meta-analysis, which were evaluated and assigned methodological validity scores by Cochrane reviewers. The scores corresponded to six validity categories, namely random sequence generation, allocation concealment, blinding of participants, staff and results analysts, incomplete reporting, and incomplete data of participants and other errors. The internal validity scores of the clinical trial articles were categorized into three levels: positive (+) indicating a low probability of bias

errors, negative (-) suggesting a high probability of bias errors, and a question mark (?) indicating uncertain bias errors. The bias error uncertainty score was assigned when the Cochrane reviewer did not obtain sufficient evidence to determine the validity of the articles. To increase the accuracy of both scores, we considered the uncertainty of bias errors and the possibility of high bias errors to be zero, while the score for a low probability of bias errors was set at 1. The resulting evaluation score represented the average score across the six categories. In the next step, the list of the clinical trial articles used in the meta-analysis of systematic review articles was downloaded and extracted from the HTML text by using the Knime software. Then, in order to increase the accuracy of identifying clinical trial articles, we only extracted articles with PubMed numbers. Additionally, bibliographic information and abstracts (abstract, title, and keywords) were extracted from the PubMed database, and their respective methodological validity scores were also incorporated by using the Python program. The 3,793 open access systematic review articles analyzed herein contained a total of 112,537 clinical trial articles used in meta-analyses, of which 52,167 had PubMed numbers. Since at least 50 documents per question were required for evaluating the effectiveness of information retrieval research, the sample needed to be refined accordingly. Consequently, 132 open access systematic review articles were identified that met two criteria: containing at least 50 clinical trial articles in their respective meta-analyses and having a PubMed ID." This refining process resulted in a final count of 9,063 clinical trial articles for analysis. To determine the standard methodological elements, we conducted a content analysis of references, standards, and educational sources pertaining to clinical trial methodology. The Google search engine was used to perform a search, which

Efectiveness of methodological elements on relevance

resulted in the retrieval of 3,903 books. The retrieved book titles were carefully examined, and 730 book titles relevant to the research topic were selected. No information regarding courses related to research methodology in medicine, critical appraisal in medicine, or evidence-based medicine was found on the websites of the world's most prestigious universities (i.e., top 100 universities on the Times ranking system list). [34] Furthermore, the websites of universities and educational centers worldwide (<https://opensyllabus.org/>) had very limited courses related to the field of medicine, with no courses specifically focused on medical research methods. However, a Google search for other websites led to the identification of training programs and courses offered by the center for evidence-based medicine center (<https://www.cebm.net/>). The books recommended for these courses were retrieved, resulting in the addition of 18 book titles to the list of evaluated books. Next, the books were downloaded in full text. Since medical texts were included in this research with the aim of enriching keywords related to bias errors in clinical trial articles, we decided to check the text, title, main topics, chapters, and back-of-the-book indexes of the downloaded books for the words "bias" and "clinical trial." From the downloaded books, 108 books were excluded because of lacking content relevant to clinical trials or bias errors, or covering irrelevant topics. Additionally, 13 books were excluded due to their unavailability in a machine-readable format. Finally, 168 books were selected for further analysis. Then, the selected books were examined in terms of their citation count, author affiliation, and publisher credibility. A total of 116 books with good citation records (at least 100 citations), written by academic authors, and published by reputable entities in the medical field were selected. To expand the basic methodological elements, we applied text preprocessing techniques to the 116

retrieved books. These techniques included converting uppercase letters to lowercase, stemming words, removing extraneous letters and punctuation marks, applying a stopword filter, and filtering out numbers and two-character meaningless words. After that, the Term Frequency-Inverse Document Frequency (TF-IDF) technique was employed to determine the weighted frequency score of each key word. Subsequently, the same pre-processing steps were performed on all the Cochrane reviewers' comments (basic methodological elements). Then, common words between the reference books and the Cochrane reviewers' comments that had more than one TF-IDF score were identified and added. In this research, the cosine similarity measure between TF-IDF values was used to determine textual similarity. To further assess the relevance of related articles, we employed the MeSH semantic similarity score, also known as semantic similarity, which is calculated by using the Jacard scale. The abstract, basic methodological elements, and standard methodological elements were represented as single words (Unigram). SPSS version 23 and the Kolmogorov–Smirnov test were used to examine the normality of data distribution. For data with a normal distribution, a paired t-test was utilized, while for data with a non-normal distribution, the Wilcoxon signed-rank test was employed.

Results

"According to the paired t-test results, there was negative evidence indicating a difference in the average NDCG score across all precision points (10th, 20th, 30th, 50th, 70th, and 100th retrieved results) for each of the four groups. "According to the paired t-test results, there was negative evidence indicating a difference in the average NDCG score across all precision points (10th, 20th, 30th, 50th, 70th, and 100th retrieved results) for each of the four groups. The difference between the

averages of the simple abstract with the basic and standard methodological elements in the abstract and the averages of the NDCG score of the simple abstract with the basic methodological elements in the abstract was significant and negative in all accuracy points. Therefore, weighting the abstract based on methodological elements reduced its ability to represent relevance.

Discussion

Methodological validity is a crucial indicator of article quality and plays a key role in systematic review research and meta-analysis. In this study, the selected articles served as substitutes for research questions, encompassing both content-related and methodology-related elements in their text. Therefore, these questions (articles text) cannot be considered to contain only content elements - and unrelated to methodology and maybe it effects the result. However, the results showed that the event or weighting of methodological elements did not help to improve the retrieval system. In addition, evaluating the validity of articles is very important in the field of evidence-based medicine. Therefore, retrieving relevant articles with higher validity can help improve the effectiveness of databases. According to the results of the present study, the occurrence or weighting of methodological elements in the abstract text has no effect on the improvement of relevance ranking or may even have a negative effect, suggesting that the validity of articles may be overlooked when they are ranked on the basis of relevance. This is contrary to the theoretical discussions that highlight validity as an indicator of relevance judgment. In practice, this indicator is not effective in ranking documents. Therefore, it is crucial to conduct further research aligned with and similar to the present study, focusing on the role of methodological sections, findings, and elements in the ranking and relevance of documents, particularly with respect to

thematic relevance. Additionally, examining integrated ranking solutions that consider both relevance and validity of articles, as well as assessing the impact of methodological elements, can contribute to enhancing the effectiveness of databases.

The results of the present study showed that textual methodological elements have either no impact or a negative impact on the improvement of relevance ranking at the abstract level of articles. This disparity can adversely affect the ranking results of systems whose users seek both relevant and valid article collections. Further research is needed to gain a more precise understanding of this effect. If the negative impact of methodological elements on thematic relevance is confirmed, it highlights the need to develop special systems. Refinement of keywords in interactive systems can help improve the efficiency and effectiveness of relational systems. According to the findings of this research, the occurrence of methodological elements in the text or weighting them can reduce the relevant results. Attention should be paid to the characteristics of MeSH as the basis (benchmark) of this research when interpreting this finding. This tool allows subject experts to identify terms related to article topics by skimming through the abstracts. [35] These titles may also include terms related to methodologies. However, since the primary focus of the subject expert's review is to determine the article's topic, the terminology associated with methodology may not receive significant attention. Therefore, it is necessary to use additional criteria, such as judgment of relevance and validation by users, to replicate this research. fore, it is necessary to repeat this research with the help of other criteria, including judgment of relevance and validation by users. In this research, the presence of words in the text of abstracts, reviewer's opinion, and medical texts was investigated at the single word level. In natural language processing, word

Efectiveness of methodological elements on relevance

relationships can indicate semantic connections. Therefore, it is necessary to investigate the effect of simultaneous occurrences of methodological elements at the two-word (Bi-grams) to multi-word (N-grams) levels. Additionally, employing advanced techniques, such as word embedding or feature selection, is necessary to provide a more comprehensive

understanding of meaning and semantic relationships. In this research, due to the lack of methodological validity scores in irrelevant articles and relevant articles excluded from the meta-analysis, the effect of the presence of methodological elements in irrelevant articles on the relevance of the results was not investigated.

References

1. Schutze H, Pedersen J. Information retrieval based on word sense. Proceedings of the Fourth Annual Symposium on Document Analysis and Information Retrieval; 1995 Apr 24-26; Las Vegas , Nevada. 1995. p. 161-176.
2. Van Nieuwenhoven CA, Buskens E, Van Tiel FH, Bonten MJ. Relationship between methodological trial quality and the effects of selective digestive decontamination on pneumonia and mortality in critically ill patients. *JAMA*. 2001;18;286(3):335-40.
3. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clin Trials*. 1996;17(1):1-12.
4. De Vries AC, Besselink MGH, Buskens E, Ridwan BU, Schipper M, Van Erpecum KJ, et al. Randomized controlled trials of antibiotic prophylaxis in severe acute pancreatitis: Relationship between methodological quality and outcome. *Pancreatology*. 2007;7(5-6):531-8.
5. Sargeant JM, Elgie R, Valcour J, Saint-Onge J, Thompson A, Marcynuk P, et al. Methodological quality and completeness of reporting in clinical trials conducted in livestock species. *Prev Vet Med*. 2009;91(2-4):107-15.
6. Xu Y, Chen Z. Relevance judgment: What do information users consider beyond topicality? *J Am Soc Inf Sci Technol*. 2006;57(7):961-73.
7. Zhao J, Kan MY, Procter PM, Zubaidah S, Yip WK, Li GM. Improving search for evidence-based practice using information extraction. *AMIA Annu Symp Proc*.2010;2010:937-41.
8. De Bruijn B, Carini S, Kiritchenko S, Martin J, Sim I. Automated information extraction of key trial design elements from clinical trial publications. *AMIA Annu Symp Proc*. 2008;2008:141-5.
9. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annu Symp Proc*. 2006;2006:359-63.
10. Boudin F, Nie JY, Bartlett JC, Grad R, Pluye P, Dawes M. Combining classifiers for robust PICO element detection. *BMC Med Inform Decis Mak*. 2010;10(1):1-6.
11. Hassanzadeh H, Groza T, Hunter J. Identifying scientific artefacts in biomedical literature: The evidence based medicine use case. *J Biomed Inform*. 2014;49:159-70.
12. Wallace BC, Kuiper J, Sharma A, Zhu MB, Marshall IJ. Extracting PICO sentences from clinical trial reports using supervised distant supervision. *J Mach Learn Res*. 2016;17:1-26.
13. Cohen AM, Smalheiser NR, McDonagh MS, Yu C, Adams CE, Davis JM, et al. Automated confidence ranked classification of randomized controlled trial articles: An aid to evidence-based medicine. *J Am Med Inform Assoc*. 2015;22(3):707-17.
14. Bekhuis T, Demner-Fushman D. Towards automating the initial screening phase of a systematic review. In: Safran C, Reti S, Marin HF, editors. *MEDINFO 2010*. IOS Press; 2010. p. 146-50.
15. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*. 2010;11:1-11.
16. Wallace BC, Noel-Storr A, Marshall IJ, Cohen AM, Smalheiser NR, Thomas J. Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. *J Am Med Inform Assoc*. 2017;24(6):1165-8.
17. Hsu W, Speier W, Taira RK. Automated extraction of reported statistical analyses: Towards a logical representation of clinical trial literature. *AMIA Annu Symp Proc*. 2012;2012:350-9.

18. Alizadeh A, Rashidi H. Intelligent extraction of important words from important parts of scientific articles. Proceedings of the Conference on Computer Engineering and Sustainable Development with a focus on computer networking, modeling and systems security; 2013 Sep 19; Mashhad: Khavaran Institute of Higher Education. 2013. p. 1-7. [In Persian]
19. Rashidi Sharifabad K, Sotoudeh H, Mirzabigi M, Fakhrahmd M. Measuring the similarity of opinions of free referees and the content of scientific articles using natural language processing. *Librarianship and Information Organization Studies*. 2020;31(2):86-103. [In Persian]
20. Paek H, Kogan Y, Thomas P, Codish S, Krauthammer M. Shallow semantic parsing of randomized controlled trial reports. *AMIA Annu Symp Proc*. 2006(2006):604-8.
21. Kiritchenko S, De Bruijn B, Carini S, Martin J, Sim I. ExaCT: Automatic extraction of clinical trial characteristics from journal publications. *BMC Med Inform Decis Mak*. 2010;10:1-17.
22. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications. *J Am Med Inform Assoc*. 2010;17(5):507-13.
23. Jain S, Peng X, Wallace BC. Detecting twitter posts with adverse drug reactions using convolutional neural networks. In: Sarker A, Gonzalez G, editors. Proceedings of the 2nd Social Media Mining for Health Research and Applications Workshop (SMM4H 2017); 2017 Nov 4; Washington, DC, United States. 2017. p. 72–75.
24. Chung GY, Coiera E. A study of structured clinical abstracts and the semantic classification of sentences. Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing; 2007 Jun 29 ; Prague, Czech Republic. United States: Association for Computational Linguistics; 2007. p. 121-8.
25. Hatami Naghani B, Abessi M. Content analysis of science paper's by using text mining. *Business Intelligence Management Studies*. 2017;5(18):137-67. [In Persian]
26. Summerscales RL, Argamon S, Bai S, Hupert J, Schwartz A. Automatic summarization of results from clinical trials. Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. 2011 Nov 12-15; Atlanta, GA, USA. IEEE; 2012. p. 372-7.
27. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. *Syst Rev*. 2014;3:1-15.
28. Dehghan M. The feasibility of presenting a quality control model for scientific articles in the publication process based on predictive indicators of scientific impact [Master's thesis]. Shiraz: Shiraz university; 2016. [In Persian]
29. Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials. *IEEE J Biomed Health Inform*. 2014;19(4):1406-12.
30. Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: Evaluation of a system for automatically assessing bias in clinical trials. *J Am Med Inform Assoc*. 2016;23(1):193-201.
31. Marshall IJ, Kuiper J, Banner E, Wallace BC. Automating biomedical evidence synthesis: RobotReviewer. *Proc Conf Assoc Comput Linguist Meet*. 2017;2017:7-12.
32. Millard LAC, Flach PA, Higgins JPT. Machine learning to assist risk-of-bias assessments in systematic reviews. *Int J Epidemiol*. 2016;45(1):266-77.
33. Asadi A, Sotoudeh H, Abbaspour J, Fakhr-Ahmad M. The potentials of cochrane reviewers' comments and citation contexts in the recognition of randomized controlled trials' texts and their main sections. *Health Information Management*. 2020;17(4):181-8. [In Persian]
34. Times Higher Education. World University Rankings [Internet]. 2018 [cited 2018 Sep 26]. Available from: <https://www.timeshighereducation.com/world-university-rankings>
35. Cool C, Belkin N, Frieder O, Kantor P. Characteristics of text affecting relevance judgments. National online meeting. 1993;14:77-77.

تأثیر عناصر روش‌شناختی مقالات کارآزمایی بالینی در اثربخشی رتبه‌بندی نتایج بر اساس ربط آنها

عادلہ اسعدی شالی^۱، هاجر ستوده^{۲*}، جواد عباس پور^۳

^۱دکتری، علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روانشناسی، دانشگاه شیراز، شیراز، ایران.

^۲استاد، دانشکده علوم تربیتی و روانشناسی، دانشگاه شیراز، شیراز، ایران.

^۳استادیار دانشکده علوم تربیتی و روانشناسی، دانشگاه شیراز، شیراز، ایران.

اطلاعات مقاله

نویسنده مسئول:

هاجر ستوده

رایانامه:

sotudeh@shirazu.ac.ir

وصول مقاله: ۱۴۰۱/۰۷/۱۰

اصلاح نهایی: ۱۴۰۱/۰۹/۲۸

پذیرش نهایی: ۱۴۰۱/۰۹/۳۰

انتشار آنلاین: ۱۴۰۲/۰۳/۲۱

واژه‌های کلیدی:

عناصر روش‌شناختی

مقالات مرور نظام‌مند کاکرین

ربط

بازیابی اطلاعات

اعتبار روش‌شناسی

نظرات ارزیابان کاکرین

پایگاه کاکرین

کارآزمایی بالینی

چکیده

مقدمه: اعتبار روش‌شناختی از جمله عوامل تأثیرگذار بر قضاوت ربط است. هدف این مطالعه، تعیین سهم و نوع تأثیر عناصر روش‌شناختی در تبیین ربط میان مدرک بوده و پرسش و تنها جنبه اثربخشی نظام بازیابی بدون ارزیابی کارایی بررسی شده است.

روش‌ها: روش پژوهش نیمه‌تجربی (پیش‌آزمون-پس‌آزمون تک‌گروهی) و جامعه پژوهش مقالات کارآزمایی بالینی موجود در فراتحلیل مقالات مرور نظام‌مند نمایه‌شده در کاکرین و نمونه‌گیری هدفمند بود. مقالات کارآزمایی بالینی موجود در فراتحلیل دارای شماره پابمد موجود در مقالات مرور نظام‌مند دسترسی آزاد پایگاه کاکرین از ابتدا تا پایان سال ۲۰۱۸ (به عنوان پرسش نظام بازیابی) با حداقل ۵۰ مقالات کارآزمایی بالینی دائلود شد. پیش‌پردازش متن و تعیین نمره «فراوانی اصطلاح- معکوس فراوانی متن» و شباهت سرعنوان‌های موضوعی پزشکی بین این مقادیر با نرم افزار نایم انجام شده و میانگین نمره «سودمندی تجمعی نزولی نرمالیزه» در گروه‌های با رتبه‌بندی مدارک در سطح چکیده به تنهایی مقایسه شد.

یافته‌ها: نتایج آزمون تی زوجی نشان داد تفاوت میانگین آندی‌سی‌جی چکیده در هر چهار گروه در تمامی نقاط دقت مدرک منفی است. بیشترین تفاوت میانگین مربوط به عناصر روش‌شناختی پایه و استاندارد در چکیده (۰/۰۶۴-) در نقطه دقت ۱۰ (دهمین مدرک در نتایج بازیابی شده)، و کمترین تفاوت میانگین مربوط عناصر روش‌شناختی پایه در چکیده (۰/۰۲۱-) در نقطه دقت ۵۰ و ۷۰ بود.

نتیجه‌گیری: یافته‌های این پژوهش نشان داد که رویداد عناصر روش‌شناختی در متن یا وزن‌دهی به آن‌ها می‌تواند نتایج مرتبط را کاهش دهد و می‌تواند بر اثربخشی نظام بازیابی بر پایه ربط و اعتبار تأثیر منفی داشته باشد.

احتیاط بیشتر ارائه می‌کنند. [۴-۸] بنابراین، اعتبار روش‌شناختی یکی از انگیزه‌های ربط مدرک از سوی کاربران محسوب می‌شود. بررسی پژوهش‌های متن‌کاوی و پردازش زبان طبیعی مرتبط با پزشکی مبتنی بر شواهد نشان داد که با توجه به نیازهای این حوزه، برخی پژوهش‌ها در زمینه دسته‌بندی اطلاعات متن مقالات براساس ساختار شکل‌دهی پرسش پژوهش به شکل بیمار، مداخله، مقایسه یا کنترل و نتیجه (Patient or Population, Intervention or Exposure, Comparison or Control, and Outcome(PICO)، [۹-۱۴] بود. گروه دیگری از پژوهش‌ها به انتخاب خودکار مقالات مرتبط [۱۵-۱۷] و قابلیت داوری‌های آزاد در بازشناخت مقالات پزشکی براساس شباهت آن‌ها، [۱۸] پرداخته بودند. استخراج هوشمند کلمات مهم از بخش‌های مهم مقالات علمی، [۲۱،۲۲-۲۵-۱۹،۱۴،۱۰] شناسایی و دسته‌بندی انواع مقالات [۱۲،۲۶] و خلاصه‌سازی متن [۳۰-۲۷،۱۳] موضوع دسته دیگری از پژوهش‌ها بود. در این پژوهش‌ها کوشیده شده تا با فنون پردازش زبان طبیعی، متن‌کاوی و روش‌های مبتنی بر ابزارهای دانشی (مانند هستی‌شناسی اصطلاحنامه و متخصص) به راهکارهایی برای این مسائل دست یابند. به طور کلی، براساس بررسی ادبیات پژوهش، پژوهشی درباره نقش عناصر درون‌متنی از جمله عناصر روش‌شناختی (کلمات مرتبط با روش‌شناسی مدارک) در تعیین ربط و اعتبار اثر مشاهده نشد. از این رو، هدف از پژوهش حاضر، سنجش اثربخشی عناصر روش‌شناختی متن در تعیین ربط آثار است که می‌تواند گامی در پر کردن شکاف دانش در این زمینه باشد. واژگان موجود در متن مقالات، منبعی اصلی برای نمایه‌سازی، بازیابی و تعیین ربط محسوب می‌شود. براساس بررسی ادبیات پژوهش، با این که واژگان مقالات می‌تواند از انواعی مانند واژگان علمی و فنی، واژگان عمومی و واژگان روش‌شناسی باشد که نقش یکسانی در تبیین ربط ندارند، [۳۱] تاکنون کمتر به تفکیک و وزن‌دهی انواع واژگان مقالات برای تعیین سنجش اثربخشی آن‌ها در تبیین ربط پرداخته شده است. رویکرد پزشکی مبتنی بر شواهد تاکید

بازیابی اطلاعات پزشکی از جنبه‌های مختلف حائز اهمیت است. یادگیری، آموزش و مرور پیشینه پژوهش، نگارش مقالات مرور نظام‌مند و فراتحلیل‌ها و پاسخ‌دهی به پرسش‌های بالینی از جمله انگیزه‌های متخصصان سلامت برای جستجوی پایگاه‌های داده است. حجم بالای انتشارات علمی موجب شده است تا پایگاه‌های داده، روش‌های ماشینی و نیمه‌ماشینی مختلفی را جهت تسهیل استخراج اطلاعات از مقالات به کار گیرند. شناسایی و دسته‌بندی مقالات براساس روش‌های پژوهش مختلف، استخراج ماشینی بخش‌های خاصی از مقالات در سیستم‌های پرسش و پاسخ پزشکی و استخراج چکیده مقالات جهت تسهیل مراحل نگارش، مطالعه و ارزیابی اعتبار مقالات بخشی از پروژه‌های در دست اجرا جهت کاهش زمان و هزینه بازیابی اطلاعات مرتبط پایگاه برای کاربران به‌شمار می‌رود. ربط (Relevance)، معیار مهمی در ارزیابی سامانه‌های بازیابی اطلاعات است و از دغدغه‌های اصلی اندیشمندان و پژوهشگران حوزه ذخیره و بازیابی اطلاعات محسوب می‌شود. [۱] ربط، میزان مرتبط بودن مدارک بازیابی شده به نیاز اطلاعاتی کاربر و مفهومی بسیار پیچیده است. قضاوت ربط از عوامل گوناگون مانند کیفیت، روزآمدی، دسترس‌پذیری، اهمیت و اثربخشی مدارک، رابط کاربری، دانش پیشین کاربر و زمان تاثیر می‌پذیرد. [۲،۳] کیفیت اطلاعات، یکی از عوامل تاثیرگذار بر قضاوت ربط است و نه تنها کاربران عمومی پایگاه، بلکه طیف گسترده‌ای از افراد از جمله داوران، پژوهشگران و کارکنان حیطه‌های مختلف به آن توجه دارند. کیفیت اطلاعات و آثار، خود از ابعادی بسیار پیچیده و در هم تنیده تشکیل شده است. در این میان، اعتبار روش‌شناختی (Methodological Validity) یکی از جنبه‌های کیفیت است که برای مقالات علمی از اهمیت ویژه‌ای برخوردار است؛ زیرا توجه به جنبه‌های اعتبار روش‌شناختی، مقاله را به واقعیتی نزدیکتر می‌کند که در پی اندازه‌گیری و سنجش آن است و می‌توان به نتایج آن اعتماد کرد. مقالات معتبر نتایج را با

دیگر، جهت گسترش کلمات روش‌شناسی موجود در متن نظرات ارزیابان کاکرین از کلمات روش‌شناسی به دست آمده از همایندهای کلمات موجود در متن نظر ارزیابان کاکرین با متون پزشکی (به طور عمده کتاب‌های روش پژوهش پزشکی) استفاده شد. در پژوهش حاضر، عناصر روش‌شناختی پایه و عناصر روش‌شناسی استاندارد در قالب چهار گروه استخراج و آزمون شد. ابتدا کلمات مشترک نظر ارزیابان کاکرین با چکیده مقالات و سپس، کلمات مشترک ارزیابان کاکرین به همراه عناصر روش‌شناسی استاندارد به دست آمد و به متن چکیده اضافه شد. دو گروه دیگر، با افزودن همه کلمات نظر ارزیابان کاکرین به متن چکیده و سپس، افزودن کلمات نظر ارزیابان کاکرین و کلمات حاصل از متون پزشکی به متن چکیده به دست آمد. در پژوهش‌های بازایی اطلاعات و ارزیابی اثربخشی رتبه‌بندی نتایج بازایی شده، وجود تعدادی پرسش، نتایج به دست آمده و میزان ربط آن‌ها به پرسش لازم است. مقالات مرور نظام‌مند در پاسخ به پرسشی بالینی نوشته می‌شوند و از این رو، هر مقاله مرور نظام‌مند یک پرسش محسوب می‌شود. از سوی دیگر، برای نگارش مقالات مرور نظام‌مند، محققان به طور گسترده در چندین پایگاه داده جستجو کرده و مقالات مرتبط با موضوع مقاله مرور نظام‌مند را انتخاب می‌کنند. از این رو، همه مقالات استفاده شده در بخش فراتحلیل مقالات مرور نظام‌مند مقالات مرتبط با سوال بالینی محسوب می‌شوند و در نتیجه، این مقالات به عنوان جامعه پژوهش حاضر در نظر گرفته شده‌اند. در این پژوهش، به منظور نمونه‌گیری جامعه پژوهش و تهیه مجموعه آزمون، ابتدا متن ۳۷۹۳ مقاله مرور نظام‌مند پایگاه کاکرین دسترسی آزاد (مقالات مروری از ابتدا تا پایان سال ۲۰۱۸) به زبان نشانه‌گذاری ابرمتنی (Hyper text markup language (HTML)) با برنامه‌ای به زبان پایتون دانلود شد. در این مقالات، فهرستی از مقالات عموماً کارآزمایی بالینی استفاده شده در فراتحلیل وجود دارد که ارزیابان کاکرین، اعتبار روش‌شناسی آن‌ها را ارزیابی و نمره‌دهی کرده‌اند. نمره‌ها به شش مقوله اعتبار (تولید توالی تصادفی، پنهان‌سازی تخصیص، کورسازی شرکت‌کننده، کارکنان و تحلیل‌گران

دارد که در فرایند جستجوی شواهد، برای تعیین نوع فرمول PICO، [۳۲] نوع روش پژوهش مورد جستجو مشخص شود. همچنین، در بسیاری از پایگاه‌های داده پزشکی، روش‌شناسی پژوهش یکی از گزینه‌های محدودکننده جستجو به شمار می‌رود. به علاوه، انواع روش‌های پژوهش، بخشی از اصطلاحات به کار رفته در سرعنوان‌های موضوعی پزشکی (Medical Subject Headings (MeSH)) و پایگاه اف هزار (F1000) را تشکیل می‌دهند. از این گذشته، اعتبار روش‌شناختی پژوهش یکی از عوامل موثر بر قضاوت ربط به‌شمار می‌رود. [۳۳] با این همه، تاکنون پژوهشی تجربی انجام نشده است که سهم و نوع تاثیر عناصر روش‌شناختی را در تبیین ربط میان مدرک و پرسش آزموده باشد. از این رو، در این پژوهش سهم و نوع تاثیر عناصر روش‌شناختی در تبیین ربط میان مدرک و پرسش در مقالات کارآزمایی بالینی به عنوان یکی از روش‌های پژوهش مهم در حوزه پزشکی آزموده شده است.

روش‌ها

پژوهش حاضر از روش شبه آزمایشی پیش‌آزمون - پس‌آزمون تک‌گروهی به منظور دستیابی به اهداف پژوهش بهره برد. متغیر مستقل شباهت متنی، متغیر وابسته ربط متخصص و متغیر به‌هنگارساز (Manipulation) شامل دو گروه از کلمات است. عناصر روش‌شناختی پایه شامل کلمات متن نظرات ارزیابان کیفی مقالات کارآزمایی بالینی وارد شده به فراتحلیل مقالات مرور نظام‌مند نمایه شده کاکرین است که در این پژوهش ارزیابان کاکرین خوانده می‌شود. ارزیابان کاکرین دلیل انتخاب نمره اعتبار روش‌شناسی را در بخش نظرات شرح داده و بدین منظور، از کلمات روش‌شناسی استفاده می‌کنند. در چندین پژوهش، کلمات نقل قول‌های داخل نظرات را منبع اصلی کلمات اعتبار روش‌شناسی در کار خود لحاظ کرده‌اند. [۳۷-۳۳] پژوهش اسعدی و همکاران [۳۸] در سال ۲۰۲۰ نشان داد که نظرات ارزیابان کاکرین علاوه بر روش پژوهش به موضوع مقاله نیز می‌پردازد. لذا در این پژوهش، کلمات نظر ارزیابان کاکرین (عناصر روش‌شناسی پایه) به عنوان منبع اصلی کلمات نشان‌دهنده اعتبار روش‌شناسی لحاظ شد. از سوی

نتایج، ارائه گزارش ناکامل و داده‌های ناکامل شرکت‌کنندگان و سایر خطاها) مربوط بوده و اعتبار درونی مقالات کارآزمایی بالینی در قالب سه سطح مثبت (+) به معنی احتمال خطای سوگیری پایین، منفی (-) به معنی احتمال خطای سوگیری بالا و نشانه پرسش (؟) به معنی نامشخص بودن میزان خطای سوگیری سنجیده شده است. نمره نامشخص بودن میزان خطای سوگیری زمانی به مقالات تعلق می‌گیرد که ارزیابان کاکرین شواهدی مبنی بر اعتبار داشتن یا نداشتن مقالات به دست نیاورند. برای افزایش دقت و با توجه به این که ارائه اطلاعات کامل در مورد روند پژوهش نشان‌دهنده دقت نویسنده و اعتبار مقاله است، هر دو نمره نامشخص بودن میزان خطای سوگیری و احتمال خطای سوگیری بالا صفر و نمره احتمال خطای سوگیری پایین یک در نظر گرفته شد. نمره به دست آمده از این ارزیابی، میانگین نمرات شش مقوله است. برخی مقالات نمره اعتبار روش‌شناختی همه مقوله‌ها را نداشتند و بنابراین، میانگین نمره مقوله‌ها برای هر مقاله نمره اعتبار روش‌شناختی آن مقاله محسوب می‌شد. همچنین، در برخی مقالات، مقوله‌های کورسازی و داده‌های ناکامل شرکت‌کنندگان بیش از یک نمره داشته و براساس هر متغیر پژوهش ارزیابی شده بودند؛ جهت یک‌دست‌سازی نمرات، نمره میانگین در این مقوله‌ها محاسبه شد. در نهایت، جهت نرمال‌سازی داده‌ها نمرات میانگین به دست آمده نهایی بر نمره بیشینه تقسیم شد و نمرات اعتبار در بازه صفر و یک قرار گرفت. در مرحله بعد، با نرم‌افزار نایم (knime) فهرست مقالات کارآزمایی بالینی مورد استفاده در فراتحلیل مقالات مرور نظام‌مند از متن HTML بارگیری و استخراج شد. سپس، جهت افزایش دقت در شناسایی مقالات کارآزمایی بالینی، تنها مقالات دارای شماره پابمد (PubMed) استخراج شدند. در ادامه با برنامه به زبان پایتون، اطلاعات کتاب‌شناختی و چکیده آن‌ها (چکیده، عنوان و کلیدواژه‌های نویسندگان) از پایگاه پابمد استخراج و نمره اعتبار روش‌شناختی آن‌ها نیز ضمیمه شد. در ۳۷۹۳ مقاله مرور نظام‌مند دسترسی آزاد، تعداد ۱۱۲۵۳۷ مقاله در فراتحلیل‌ها به خدمت گرفته شده بودند که از این میان

۵۲۱۶۷ مقاله کارآزمایی بالینی دارای شماره پابمد بود. از آنجا که برای ارزیابی اثربخشی رتبه‌بندی بازایی اطلاعات، به حداقل ۵۰ مدرک برای هر پرسش نیاز بود نمونه می‌بایست به این لحاظ هم پالایش شود. بدین منظور، تعداد ۱۳۲ مقاله مرور نظام‌مند دسترسی آزاد شناسایی شد که حداقل ۵۰ مقاله کارآزمایی بالینی در فراتحلیل داشتند و دارای شناسه پابمد نیز بودند. به این ترتیب، تعداد مقالات کارآزمایی بالینی به ۹۰۶۳ مقاله رسید. از آنجا که هر مقاله مروری مرتبط با یک پرسش بالینی است، این عدد شمار پرسش‌های این پژوهش را نیز تعیین کرد. از میان مقالات کارآزمایی بالینی وارد شده به فراتحلیل، یک مقاله به طور تصادفی به عنوان مدرک پایه (جانشین پرسش) انتخاب شد و دیگر مقالات وارد شده به فراتحلیل‌ها به عنوان مدارک مجموعه انتخاب شدند. در این پژوهش، عبارت اعتبار روش‌شناختی تنها ناظر بر نوع اعتبار درونی روش‌شناختی است که ارزیابان کاکرین آن را سنجیده و نظر داور در متن مقالات مروری برای مقالات استفاده شده در فراتحلیل براساس مقوله‌های ابزار خطای سوگیری کاکرین (Cochrane risk of bias tool) (تولید توالی تصادفی، پنهان‌سازی تخصیص، کورسازی شرکت‌کننده، کارکنان و تحلیلگران نتایج، ارائه گزارش ناکامل و داده‌های ناکامل شرکت‌کنندگان و سایر خطاها) آمده است. عناصر روش‌شناختی پایه، واژه‌های موجود در متن نظر ارزیابان کاکرین است و منظور از عناصر روش‌شناختی استاندارد، واژه‌هایی است که در متون پزشکی مرتبط با مقالات کارآزمایی بالینی و روش‌شناسی آن استفاده شده است. این عناصر با تحلیل محتوای بخش‌های مرتبط با کارآزمایی بالینی در متون روش‌شناسی پژوهش در پزشکی، ارزیابی نقادانه در پزشکی، پزشکی مبتنی بر شواهد به دست آمد. جهت تعیین عناصر روش‌شناختی استاندارد، از تحلیل محتوای منابع مرجع، استاندارد یا آموزشی درباره روش‌شناسی کارآزمایی بالینی استفاده شد. بدین منظور، در موتور کاوش گوگل کلیدواژه‌های جدول ۱ جستجو شده و ۳۹۰۳ کتاب بازایی شد. عناوین کتب بازایی شده به دقت بررسی و ۷۳۰ عنوان

بارگیری شد. با توجه به اینکه هدف از افزودن متون پزشکی در این پژوهش غنی کردن کلیدواژه‌های مرتبط با خطای سوگیری مقالات کارآزمایی بالینی بود، لذا متن کتاب‌های بارگیری شده به لحاظ حضور کلمات سوگیری و کارآزمایی بالینی در عنوان کتاب، موضوع اصلی، فصل، نمایه آخر و داخل متن کتاب بررسی شد. از کتاب‌های بارگیری شده، ۱۰۸ کتاب از مجموعه حذف شدند؛ زیرا فاقد محتوا در زمینه کارآزمایی بالینی و یا خطای سوگیری بودند و یا موضوعی غیر مرتبط به این زمینه داشتند. فرمت ۱۳ کتاب نیز قابل خواندن با ماشین نبود و از مجموعه مرجع کنار گذاشته شد. در نهایت، ۱۶۸ کتاب در این مرحله انتخاب شد. سپس، کتاب‌ها به لحاظ میزان استناد، وابستگی دانشگاهی نویسندگان کتاب و اعتبار ناشر در حیطه پزشکی بررسی و ۱۱۶ کتاب با نویسندگان دانشگاهی دارای استناد خوب (حداقل ۱۰۰ استناد) و داشتن ناشران معروف در حوزه پزشکی انتخاب شد.

کتاب مرتبط با موضوع انتخاب شد. با بررسی وبگاه دانشگاه‌های معتبر جهان (۱۰۰ دانشگاه برتر فهرست نظام رتبه‌بندی تایمز) [۳۹] در زمینه برگزاری دوره‌های مرتبط با روش‌شناسی پژوهش در پزشکی، ارزیابی نقادانه در پزشکی، پزشکی مبتنی بر شواهد اطلاعاتی، اطلاعاتی در این زمینه به دست نیامد. دوره‌های مربوط به حیطه پزشکی نیز در وبگاه برنامه‌های درسی دانشگاه‌ها و مراکز آموزشی دنیا (<https://opensyllabus.org/>) بسیار محدود بود و دوره‌ای در زمینه روش پژوهش پزشکی در این وبگاه‌ها وجود نداشت. با جستجو در موتور کاوش گوگل در زمینه سایر وبگاه‌ها، برنامه‌های آموزشی مرکز پزشکی مبتنی بر شواهد (<https://www.cebm.net/>) و دوره‌های برگزار شده در این موسسه به همراه کتاب‌های معرفی شده برای این دوره‌ها بازیابی شد و با این روش، ۱۸ عنوان کتاب به فهرست کتاب‌های مورد ارزیابی اضافه شد. در مرحله بعد، کتاب‌ها در پایگاه‌های اشتراکی و پایگاه‌های برخط کتاب دانشگاه شیراز جستجو، ۲۸۴ عنوان کتاب بازیابی و متن کامل آن‌ها

جدول ۱: نتایج جستجوی کلیدواژه‌های مختلف و تعداد کتاب‌های بازیابی شده در گوگل بوک

ردیف	کلیدواژه جستجو	تعداد کتاب‌های بازیابی شده
۱	Clinical Research	۲۸۶
۲	Clinical Trial	۲۷۱
۳	Evidence Based Medicine	۱۷۴
۴	Evidence Based Practice	۲۸۴
۵	Critical Appraisal	۱۴۴
۶	Clinical Trials	۲۸۱
۷	Epidemiology	۲۴۹
۸	Randomized Trial	۲۳۲
۹	Randomized Trials	۱۱۶
۱۰	Randomized Trial	۲۸۹
۱۱	Randomized Trials	۲۰۷
۱۲	Trials	۲۸۵
۱۳	Trial	۲۸۵
۱۴	Bias	۳۸۹
۱۵	Risk of Bias	۱۷
۱۶	RCT	۷۹
۱۷	Medical Research	۳۱۵
	تعداد کل جستجو شده در بخش کتاب‌های گوگل	۳۹۰۳

برای درجه بندی بیشتر ربط مقالات مرتبط، از نمره شباهت معنایی MeSH یا به اختصار شباهت معنایی استفاده شد که با سنجہ ژاکارد (Jacard) محاسبه می شود. این سنجہ، شباهت میان دو متن را از تقسیم اشتراک اصطلاحات آن دو بر اجتماع اصطلاحات اندازه می گیرد. در این پژوهش، شباهت معنایی میان دو متن، براساس شباهت سرعنوان های MeSH تخصیص یافته به مدارک جانشین پرسش و همچنین، مدارک مجموعه محاسبه شد و به عنوان نمره ربط متخصص جهت رتبه بندی مدارک براساس ربط آن ها در نظر گرفته شد. در مرحله پیش پردازش متن، عنوان، چکیده و کلیدواژه ها به عنوان بازنمون های متن با هم ترکیب شدند. جهت پیش پردازش متن در برنامه نایم از گره های حذف حروف و علائم سجاوندی زاید، حذف کلمات ایستا، حذف اعداد، حذف کلمات بی معنی دو حرفی، تبدیل حروف بزرگ به کوچک و استخراج ریشه کلمات استفاده شد. چکیده، عناصر روش شناختی پایه و عناصر روش شناختی استاندارد به صورت تک واژه (Unigram) درآمد. نرمال بودن توزیع داده ها با نرم افزار SPSS نسخه ۲۳ و آزمون کولموگوروف اسمیرنوف (Kolmogorov-Smirnov) بررسی شد. در مواردی با توزیع نرمال از آزمون تی جفتی (Paired T-test) و در مواردی با توزیع غیرنرمال از آزمون ویلکاکسون (-Wilcoxon Signed Rank Test) استفاده شد (جدول پیوست یک). تفاوت میانگین سودمندی تجمعی نزولی نرمالیزه (Discounted Cumulative Gain (nDCG) (طبق رابطه یک) چکیده، معیاری برای ارزیابی کیفیت رتبه بندی مدارک ارزیابی شده است. این معیار در پژوهش حاضر، به مقایسه رتبه بندی نتایج قبل و بعد از افزودن عناصر روش شناسی به متن چکیده می پردازد و با آن، میزان تاثیر عناصر روش شناسی بر نزدیک شدن رتبه بندی مدارک در سطح چکیده به رتبه بندی متخصص ارزیابی شد.

$$DCG_p = \sum_{i=1}^p \frac{rel_i}{\log_2(i+1)}$$

$$IDCG_p = \sum_{i=1}^{|REL_p|} \frac{rel_i}{\log_2(i+1)}$$

جهت گسترش عناصر روش شناختی پایه، در ۱۱۶ کتاب بازیابی شده، پیش پردازش متن (Text preprocessing) شامل: تبدیل حروف بزرگ به کوچک (Uppercase to lowercase)، ریشه یابی واژه ها (stemming) و حذف حروف زاید، نشانه های سجاوندی (punctuation erasure)، ناواژه ها (Stopword filter)، اعداد و واژه های بی معنی دو حرفی (2-character filter) انجام شد. پس از آن برای هر واژه، نمره فراوانی وزنی کلمات کلیدی مدارک فراوانی اصطلاح - معکوس فراوانی متن (Term Frequency - Inverse Document Frequency (TF-IDF)) به عنوان وزن هر کلمه تعیین شد. در ادامه، همین مراحل پیش پردازش بر روی همه نظرات ارزیابان کاکرین (عناصر روش شناختی پایه) نیز انجام شد. سپس، کلمات مشترک موجود در متن نظرات ارزیابان کاکرین با متون پزشکی و کلمات همآیند آن ها در متون پزشکی شناسایی و اضافه شد. به دلیل حجم بالای داده ها و محدودیت سخت افزاری، تنها کلماتی انتخاب شد که نمره تی اف آی دی اف آن ها بیشتر از یک بود. همچنین، برخلاف کلمات نظر ارزیابان کاکرین، عناصر روش شناختی استاندارد فاقد نمره اعتبار روش شناختی بود. به همین دلیل و برای بهبود دقت، عناصر روش شناختی استاندارد به عنوان همآیند انتخاب شدند که تنها در یکی از دسته های خطر سوگیری بالا و یا خطر سوگیری پایین قرار داشتند. در این پژوهش جهت تعیین شباهت متنی، از سنجہ شباهت کسینوسی بین مقادیر تی اف آی دی اف استفاده شد. این سنجہ، مقادیر صفر تا یک را در بر می گیرد که صفر به معنای نداشتن شباهت و یک به معنای شباهت کامل است. شایان ذکر است از شباهت متنی در متون معادل ربط سامانه ای یا ربط الگوریتمی یاد می شود. [۴۰] ملاک ربط پایه در این پژوهش شباهت متنی است. همچنین، برای محاسبه ربط متخصص، همه مقالات کارآزمایی بالینی وارد شده به فراتحلیل به عنوان مقالات مرتبط در نظر گرفته شد؛ زیرا متخصصان کاکرین پس از تایید ربط مقاله با موضوع مورد جستجو، اقدام به ارزیابی آن می کنند. به این ترتیب، فهرست مقالاتی آورده شده در مقالات مروری، مرتبط تلقی می شوند.

بازیابی شده) و کمترین تفاوت میانگین مربوط به عناصر روش شناختی پایه در چکیده (۰/۲۱-) در نقطه دقت ۵۰ و ۷۰ بود. تفاوت میان میانگین‌های چکیده ساده با عناصر روش شناختی پایه و استاندارد در چکیده و میانگین‌های نمره ان دی سی جی چکیده ساده با عناصر روش شناختی پایه در چکیده در همه نقاط دقت معنی‌دار و منفی بود. بنابراین، وزن-دهی چکیده براساس عناصر روش شناختی، توان آن را در بازنمایی ربط کاهش داد. جدول سه نتایج را برای گسترش چکیده با عناصر روش شناختی پایه و استاندارد به تصویر می‌کشد. آزمایش‌ها برای سطوح دیگر مانند عناصر روش شناختی پایه در چکیده، عناصر روش شناختی استاندارد در چکیده و عناصر روش شناختی پایه و استاندارد در چکیده نیز انجام شد.

$$nDCG_p = \frac{DCG_p}{IDCG_p}$$

رابطه ۱: فرمول نمره ان دی سی جی

یافته‌ها

بنابر نتایج آزمون تی زوجی، تفاوت میانگین نمره ان دی سی جی (معیاری جهت ارزیابی اثربخشی نتایج جستجو در پایگاه داده است که مقداری عددی بین صفر و یک دارد و صحت رتبه‌بندی نتایج را اندازه‌گیری می‌کند) چکیده در هر چهار گروه و همه نقاط دقت (۱۰، ۲۰، ۳۰، ۵۰، ۷۰ و ۱۰۰) آمین نتیجه بازیابی شده) مدرک منفی است. بیشترین تفاوت میانگین مربوط به عناصر روش شناختی پایه و استاندارد در چکیده (۰/۶۴-) در نقطه دقت ۱۰ (دهمین مدرک در نتایج

جدول ۳: مقایسه اثربخشی چکیده قبل و بعد از گسترش با عناصر روش شناختی پایه و استاندارد

نقطه دقت	متن آزمون شده	اختلاف میانگین با حالت ساده	انحراف استاندارد	میانگین خطای استاندارد	فاصله اطمینان ۹۵٪		آماره تی / زد	معنی داری
					حد بالا	حد پایین		
۱۰	چکیده به همراه عناصر روش شناختی پایه	-۰/۰۵	۰/۱۵۸۰	۰/۱۴۰	-۰/۰۷۸	-۰/۰۲۳	-۳/۶۰۸	۰
	چکیده به همراه عناصر روش شناختی پایه و استاندارد	-۰/۰۶۲	۰/۱۴۸۰	۰/۱۳۰	-۰/۰۸۸	-۰/۰۳۶	-۴/۷۵۸	۰
	عناصر روش شناختی پایه در چکیده	-۰/۰۴	۰/۱۰۵	۰/۰۰۹	-۰/۰۵۸	-۰/۰۲۱	-۴/۲۸۸	۰
	عناصر روش شناختی پایه و استاندارد در چکیده	-۰/۰۶۴	۰/۱۲۸	۰/۰۱۲	-۰/۰۸۸	-۰/۰۴	-۵/۲۵۲	۰
۲۰	چکیده به همراه عناصر روش شناختی پایه	-۰/۰۴	۰/۱۱۶	۰/۰۱	-۰/۰۶۱	-۰/۰۲	-۳/۹۲۱	۰
	چکیده به همراه عناصر روش شناختی پایه و استاندارد	-۰/۰۵	۰/۱۰۳	۰/۰۰۹	-۰/۰۶۸	-۰/۰۳۱	-۵/۴۰۸	۰
	عناصر روش شناختی پایه در چکیده	-۰/۰۳۱	۰/۰۶۷	۰/۰۰۶	-۰/۰۴۳	-۰/۰۱۹	-۵/۱۷۸	۰
	عناصر روش شناختی پایه و استاندارد در چکیده	-۰/۰۵	۰/۰۹۴	۰/۰۰۸	-۰/۰۶۷	-۰/۰۳۴	-۵/۹۸	۰
۳۰	چکیده به همراه عناصر روش شناختی پایه	-۰/۰۳۷	۰/۰۹۵	۰/۰۰۹	-۰/۰۵۴	-۰/۰۲	-۴/۳۱۲	۰
	چکیده به همراه عناصر روش شناختی پایه و استاندارد	-۰/۰۴۴	۰/۰۸۸	۰/۰۰۸	-۰/۰۶	-۰/۰۲۸	-۵/۵۳۹	۰
	عناصر روش شناختی پایه در چکیده	-۰/۰۲۲	۰/۰۵۷	۰/۰۰۵	-۰/۰۳۲	-۰/۰۱۱	-۴/۲۰۵	۰
	عناصر روش شناختی پایه و استاندارد در چکیده	-۰/۰۴۲	۰/۰۸۱	۰/۰۰۷	-۰/۰۵۷	-۰/۰۲۸	-۵/۷۷	۰
۵۰	چکیده به همراه عناصر روش شناختی پایه	-۰/۰۳۹	۰/۰۶۷	۰/۰۰۸	-۰/۰۵۴	-۰/۰۲۳	-۴/۹۷۶	۰
	چکیده به همراه عناصر روش شناختی پایه و استاندارد	-۰/۰۴۶	۰/۰۶۷	۰/۰۰۸	-۰/۰۶۱	-۰/۰۳۱	-۵/۹۶۱	۰
	عناصر روش شناختی پایه در چکیده	-۰/۰۲۱	۰/۰۴۳	۰/۰۰۵	-۰/۰۳۱	-۰/۰۱۱	-۴/۲۶۸	۰
	عناصر روش شناختی پایه و استاندارد در چکیده	-۰/۰۴	۰/۰۵۴	۰/۰۰۶	-۰/۰۵۳	-۰/۰۲۸	-۶/۴۹۸	۰
۷۰	چکیده به همراه عناصر روش شناختی پایه	-۰/۰۴۴	۰/۰۵۸	۰/۰۱۱	-۰/۰۶۶	-۰/۰۲۲	-۴/۰۴۸	۰
	چکیده به همراه عناصر روش شناختی پایه و استاندارد	-۰/۰۵۳	۰/۰۶۳	۰/۰۱۲	-۰/۰۷۷	-۰/۰۲۹	-۴/۵۵۴	۰
	عناصر روش شناختی پایه در چکیده	-۰/۰۲۱	۰/۰۴۳	۰/۰۰۸	-۰/۰۳۸	-۰/۰۰۵	-۲/۶۷۱	۰/۰۱۲
	عناصر روش شناختی پایه و استاندارد در چکیده	-۰/۰۴۱	۰/۰۴۹	۰/۰۰۹	-۰/۰۶	-۰/۰۲۲	-۴/۴۸۱	۰

جدول ۳: ادامه

معنی داری	آماره تی / زد	فاصله اطمینان ۹۵٪		میانگین خطای استاندارد	انحراف استاندارد	اختلاف میانگین با حالت ساده	متن آزمون شده	نقطه دقت
		حد بالا	حد پایین					
۰/۰۰۱	-۴/۰۴۴	-۰/۰۲۲	-۰/۰۷۲	۰/۰۱۲	۰/۰۴۵	-۰/۰۴۷	چکیده به همراه عناصر روش شناختی پایه	۱۰۰
۰	-۴/۷۹۱	-۰/۰۳	-۰/۰۸	۰/۰۱۱	۰/۰۴۵	-۰/۰۵۵	چکیده به همراه عناصر روش شناختی پایه و استاندارد	
۰/۰۲۴	-۲/۵۲۳	-۰/۰۰۴	-۰/۰۴۸	۰/۰۱	۰/۰۴	-۰/۰۲۶	عناصر روش شناختی پایه در چکیده	
۰	-۴/۷۲۹	-۰/۰۲	-۰/۰۵۴	۰/۰۰۸	۰/۰۳۱	-۰/۰۳۷	عناصر روش شناختی پایه و استاندارد در چکیده	

نتیجه گیری

ارزیابی های خود از نتایج بازیابی نه تنها به ربط بلکه به اعتبار منابع نیز توجه دارند. [۴۳] در تفسیر این یافته باید به ویژگی های MeSH به عنوان مبنای (Benchmark) این پژوهش توجه داشت. متخصصان موضوعی با مطالعه چکیده مقالات و با این ابزار، اصطلاحات مرتبط با موضوع مقاله [۴۴] را تعیین می کنند. در این سرعنوان ها، اصطلاحات مرتبط با روش شناسی ها نیز به کار رفته است. با این حال، با توجه به هدف تعیین موضوع در بررسی مقاله به دست متخصص موضوعی، ممکن است به اصطلاحات مرتبط با روش شناسی توجه زیادی نشود. بنابراین، تکرار این پژوهش به کمک معیارهای دیگر، از جمله قضاوت ربط و اعتبارسنجی از سوی کاربران ضروری است. اعتبار روش شناختی یکی از شاخص های مهم کیفیت مقالات به شمار می رود. ارزیابی اعتبار مقالات یکی از مراحل اجرای پژوهش های مرور نظام مند و فراتحلیل ها است. به علاوه، ارزیابی اعتبار مقالات در حوزه پزشکی مبتنی بر شواهد از اهمیت بالایی برخوردار است. بنابراین، بازیابی مقالات مرتبط با اعتبار بالاتر می تواند بر بهبود اثر بخشی پایگاه های داده کمک کند. براساس نتایج پژوهش حاضر، رویداد یا وزن دهی عناصر روش شناختی در متن چکیده بر بهبود رتبه بندی ربط بی تاثیر است یا تاثیر منفی دارد که نشان می دهد اعتبار مقالات ممکن است در رتبه بندی آن ها براساس ربط نادیده گرفته شود و برخلاف مباحث نظری رشته که اعتبار یکی از شاخص های قضاوت ربط است، در عمل این شاخص در رتبه بندی مدارک موثر نباشد. لذا انجام و تکرار پژوهش های هم راستا و مشابه با پژوهش حاضر

بررسی سهم عناصر روش شناختی در تبیین ربط مدارک از اهمیت به سزایی برخوردار است؛ زیرا هر چند اجزای مختلف یک مدرک می توانند به نوعی در تبیین محتوای آن موثر و با موضوع اثر مرتبط باشند اما شدت این ربط برای همه اجزا برابر نیست. به علاوه، اگرچه پرسش کاربران می تواند پیرامون هر جزء از مدرک از جمله روش شناسی آن باشد و در پزشکی مبتنی بر شواهد نیز بر ضرورت توجه به روش شناسی در تدوین راهبرد جستجو تاکید شده است. [۴۱] اما ممکن است پرسش های کاربران بیش از روش شناسی معطوف به مسئله و موضوع مقالات [۴۲] یا نتایج و راهکارهای آن ها باشند. [۹] در این صورت، وجود عناصر روش شناسی در چکیده کمکی به تبیین ربط نخواهد کرد. بنا بر یافته های این پژوهش، عناصر روش شناختی به طور مستقل یا برای گسترش چکیده، بر رتبه بندی نتایج ربط تاثیری ندارد یا دارای تاثیر منفی است. به عبارت دیگر، رویداد عناصر روش شناختی در متن یا وزن دهی به آن ها می تواند نتایج مرتبط را کاهش دهد. آشکار است که در این پژوهش، مقالات پایه به عنوان جانشین پرسش انتخاب شدند که در آن ها علاوه بر عناصر مرتبط با محتوا، عناصر مرتبط با روش شناسی نیز دیده می شود. بنابراین، این پرسش ها را نمی توان تنها حاوی عناصر محتوایی و نامرتبط با روش شناسی دانست. با این حال، نتایج نشان داد که رویداد یا وزن دهی به عناصر روش شناختی کمکی به بهبود ربط نمی کند. اهمیت این یافته ها در دانستن این امر است که کاربران در

ارزیابی ربط استفاده شود. از آنجا که انتخاب کلیدواژه‌های MeSH به دست متخصصان موضوعی با مطالعه چکیده و با هدف تعیین موضوع مقالات است، احتمال دارد تمرکز آن‌ها بر ربط موضوعی بیشتر از ویژگی‌های روش‌شناختی باشد. به علاوه، با وجود توجه به روش‌شناسی‌ها در MeSH، جزئیات روش‌شناسی‌ها از جمله جزئیات مرتبط با اعتبار آن‌ها پوشش داده نمی‌شود. از این رو، نیاز به پژوهش‌های بیشتر با ابزارهای متعدد و جامع وجود دارد. در این پژوهش، حضور کلمات در متن چکیده، نظر داور و متون پزشکی در سطح تک‌واژه بررسی شد. در زبان طبیعی ارتباط میان واژگان می‌تواند مبین روابط معنایی باشد؛ بنابراین، ضروری است تاثیر حضور همزمان عناصر روش‌شناختی در سطح دوواژه‌ای (Bi-grams) تا چندواژه‌ای (N-grams) نیز بررسی شود. همچنین، استفاده از فنون پیشرفته‌تر برای تبیین بهتر معنا و روابط معنایی مانند تعبیه واژگان (Embedding word) یا انتخاب ویژگی (Feature selection) نیز پیشنهاد می‌شود. در این پژوهش، به دلیل نبود نمره اعتبار روش‌شناختی در مقالات نامرتبط و مقالات مرتبط کنار گذاشته از فراتحلیل، تاثیر حضور عناصر روش‌شناختی در مقالات نامرتبط بر ربط نتایج بررسی نشد.

ملاحظات اخلاقی

رعایت دستورالعمل‌های اخلاقی: این مقاله حاصل بخشی از پایان‌نامه با عنوان «تاثیر عناصر روش‌شناختی متنی بر اثربخشی رتبه‌بندی نتایج بازیابی بر پایه ربط و اعتبار روش‌شناختی مقالات کارآزمایی بالینی» در مقطع دکتری، مصوب دانشگاه شیراز در سال ۱۳۹۸ است.

حمایت مالی: این پژوهش با حمایت مالی دانشگاه شیراز، ایران انجام شده است.

تضاد منافع: نویسندگان اظهار داشتند که تضاد منافی وجود ندارد.

تشکر و قدردانی: در اینجا لازم می‌دانیم از جناب آقای دکتر فرهاد شکرانه، نماینده کاکرین در ایران و کلیه افرادی که در این پژوهش ما را یاری نمودند تشکر و قدردانی نماییم.

و در زمینه نقش بخش‌های روش‌شناسی، یافته‌ها و عناصر روش‌شناختی در رتبه‌بندی و ربط (به ویژه ربط موضوعی) مدارک حائز اهمیت است. همچنین، بررسی راهکارهای رتبه‌بندی تلفیقی ربط و اعتبار مقالات و تاثیر عناصر روش‌شناختی بر آن، می‌تواند به بهبود اثربخشی پایگاه‌های داده کمک کند. نتایج پژوهش حاضر نشان داد که عناصر روش‌شناختی متنی بر بهبود رتبه‌بندی ربط در سطح چکیده مقالات بی‌تاثیر هستند یا تاثیر منفی دارند. این ناهمسویی می‌تواند بر نتایج رتبه‌بندی سامانه‌هایی تاثیر نامطلوب داشته باشد که کاربران آن‌ها به دنبال دستیابی به مجموعه مقالات مرتبط و در عین حال معتبر هستند. پژوهش‌های بیشتر به منظور شناسایی دقیق‌تر این تاثیر ضروری به نظر می‌رسد. اگر تاثیر منفی عناصر روش‌شناختی بر ربط موضوعی تحکیم شود، لزوم توسعه سامانه‌های بازیابی هدفمند را برجسته می‌سازد. پالایش هدفمند کلیدواژه‌ها در سامانه‌های تعاملی می‌تواند به بهبود کارایی و اثربخشی نظام‌های ربط کمک کند. با توجه به نبود پژوهش مشابه در زمینه میزان تاثیر عناصر روش‌شناسی در بهبود ربط، امکان مقایسه نتایج این پژوهش با پژوهش‌های دیگر مقدور نبود. با این وجود، پژوهش حاضر را می‌توان در زمره پژوهش‌های مرتبط با حوزه شبه بازخورد ربط یا بازخورد ربط کور- (Blind or pseudo-relevance feedback) و همچنین، حوزه آموزش از راه دور در نظر گرفت. در شبه بازخورد، تلاش بر آن است تا مرتبط‌ترین آثار از میان نتایج بازیابی اولیه شناسایی و با پردازش زبان طبیعی، مهم‌ترین واژگان آن‌ها به پرسش افزوده شود. سپس، بار دیگر پرسش گسترش یافته به سامانه داده شده و نتایج حاصل از آن در اختیار کاربر قرار گیرد. [۴۵] بر این اساس، نتایج پژوهش حاضر کاربردهایی برای بهبود سامانه‌های بازخورد ربط کور دارد. اجرای این پژوهش با محدودیت‌ها و دشواری‌هایی همراه بود. نبود استاندارد طلایی برای ارزیابی ربط و اعتبار مقالات و همچنین، دشواری تعیین ربط از نظر کاربر واقعی باعث شد تا در این پژوهش از معیار شباهت معنایی اصطلاحات MeSH برای

References

1. Fatahi R. Analysis of factors affecting the relativity of relevance in information retrieval systems. *Informatics*. 2004;2(1):7-22. [In Persian].
2. Khalowi M. Its relevance and meaning in information retrieval. *Iranian Journal of Information Processing & Management*. 2008;23(3):105-18. [In Persian].
3. Goffman W, Newill VA. Communication and epidemic processes. *Proc R Soc Lond A Math Phys Sci*. 1967;298(1454): 316-34.
4. Schutze H, Pedersen J. Information retrieval based on word sense. *Proceedings of the Fourth Annual Symposium on Document Analysis and Information Retrieval*; 1995 Apr 24-26; Las Vegas , Nevada. 1995. p. 161-176.
5. Van Nieuwenhoven CA, Buskens E, Van Tiel FH, Bonten MJ. Relationship between methodological trial quality and the effects of selective digestive decontamination on pneumonia and mortality in critically ill patients. *JAMA*. 2001;18;286(3):335-40.
6. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clin Trials*. 1996;17(1):1-12.
7. De Vries AC, Besselink MGH, Buskens E, Ridwan BU, Schipper M, Van Erpecum KJ, et al. Randomized controlled trials of antibiotic prophylaxis in severe acute pancreatitis: Relationship between methodological quality and outcome. *Pancreatolgy*. 2007;7(5-6):531-8.
8. Sargeant JM, Elgie R, Valcour J, Saint-Onge J, Thompson A, Marcynuk P, et al. Methodological quality and completeness of reporting in clinical trials conducted in livestock species. *Prev Vet Med*. 2009;91(2-4):107-15.
9. Xu Y, Chen Z. Relevance judgment: What do information users consider beyond topicality? *J Am Soc Inf Sci Technol*. 2006;57(7):961-73.
10. Zhao J, Kan MY, Procter PM, Zubaidah S, Yip WK, Li GM. Improving search for evidence-based practice using information extraction. *AMIA Annu Symp Proc*. 2010;2010:937-41.
11. De Bruijn B, Carini S, Kiritchenko S, Martin J, Sim I. Automated information extraction of key trial design elements from clinical trial publications. *AMIA Annu Symp Proc*. 2008;2008:141-5.
12. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annu Symp Proc*. 2006;2006:359-63.
13. Boudin F, Nie JY, Bartlett JC, Grad R, Pluye P, Dawes M. Combining classifiers for robust PICO element detection. *BMC Med Inform Decis Mak*. 2010;10(1):1-6.
14. Hassanzadeh H, Groza T, Hunter J. Identifying scientific artefacts in biomedical literature: The evidence based medicine use case. *J Biomed Inform*. 2014;49:159-70.
15. Wallace BC, Kuiper J, Sharma A, Zhu MB, Marshall IJ. Extracting PICO sentences from clinical trial reports using supervised distant supervision. *J Mach Learn Res*. 2016;17:1-26.
16. Cohen AM, Smalheiser NR, McDonagh MS, Yu C, Adams CE, Davis JM, et al. Automated confidence ranked classification of randomized controlled trial articles: An aid to evidence-based medicine. *J Am Med Inform Assoc*. 2015;22(3):707-17.
17. Bekhuis T, Demner-Fushman D. Towards automating the initial screening phase of a systematic review. In: Safran C, Reti S, Marin HF, editors. *MEDINFO 2010*. IOS Press; 2010. p. 146-50.
18. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*. 2010;11:1-11.
19. Wallace BC, Noel-Storr A, Marshall IJ, Cohen AM, Smalheiser NR, Thomas J. Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. *J Am Med Inform Assoc*. 2017;24(6):1165-8.

20. Hsu W, Speier W, Taira RK. Automated extraction of reported statistical analyses: Towards a logical representation of clinical trial literature. *AMIA Annu Symp Proc.* 2012;2012:350-9.
21. Alizadeh A, Rashidi H. Intelligent extraction of important words from important parts of scientific articles. *Proceedings of the Conference on Computer Engineering and Sustainable Development with a focus on computer networking, modeling and systems security; 2013 Sep 19; Mashhad: Khavaran Institute of Higher Education.* 2013. p. 1-7. [In Persian]
22. Rashidi Sharifabad K, Sotoudeh H, Mirzabigi M, Fakhrahmd M. Measuring the similarity of opinions of free referees and the content of scientific articles using natural language processing. *Librarianship and Information Organization Studies.* 2020;31(2):86-103. [In Persian]
23. Paek H, Kogan Y, Thomas P, Codish S, Krauthammer M. Shallow semantic parsing of randomized controlled trial reports. *AMIA Annu Symp Proc.* 2006(2006):604-8.
24. Kiritchenko S, De Bruijn B, Carini S, Martin J, Sim I. ExaCT: Automatic extraction of clinical trial characteristics from journal publications. *BMC Med Inform Decis Mak.* 2010;10:1-17.
25. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications. *J Am Med Inform Assoc.* 2010;17(5):507-13.
26. Jain S, Peng X, Wallace BC. Detecting twitter posts with adverse drug reactions using convolutional neural networks. In: Sarker A, Gonzalez G, editors. *Proceedings of the 2nd Social Media Mining for Health Research and Applications Workshop (SMM4H 2017); 2017 Nov 4; Washington, DC, United States.* 2017. p. 72-75.
27. Chung GY, Coiera E. A study of structured clinical abstracts and the semantic classification of sentences. *Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing; 2007 Jun 29; Prague, Czech Republic. United States: Association for Computational Linguistics; 2007.* p. 121-8.
28. Hatami Naghani B, Abessi M. Content analysis of science paper's by using text mining. *Business Intelligence Management Studies.* 2017;5(18):137-67. [In Persian]
29. Summerscales RL, Argamon S, Bai S, Hupert J, Schwartz A. Automatic summarization of results from clinical trials. *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine.* 2011 Nov 12-15; Atlanta, GA, USA. IEEE; 2012. p. 372-7.
30. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. *Syst Rev.* 2014;3:1-15.
31. Yoon SH, Kim SW, Park S. C-Rank: A link-based similarity measure for scientific literature databases. *Inf Sci.* 2016;326:25-40.
32. Heneghan C, Badenoch D. *Evidence-based medicine toolkit.* 2nd ed. John Wiley & Sons; 2013.
33. Dehghan M. The feasibility of presenting a quality control model for scientific articles in the publication process based on predictive indicators of scientific impact [Master's thesis]. Shiraz: Shiraz university; 2016. [In Persian]
34. Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials. *IEEE J Biomed Health Inform.* 2014;19(4):1406-12.
35. Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: Evaluation of a system for automatically assessing bias in clinical trials. *J Am Med Inform Assoc.* 2016;23(1):193-201.
36. Marshall IJ, Kuiper J, Banner E, Wallace BC. Automating biomedical evidence synthesis: RobotReviewer. *Proc Conf Assoc Comput Linguist Meet.* 2017;2017:7-12.
37. Millard LAC, Flach PA, Higgins JPT. Machine learning to assist risk-of-bias assessments in systematic reviews. *Int J Epidemiol.* 2016;45(1):266-77.

38. Asadi A, Sotudeh H, Abbaspour J, Fakhr-Ahmad M. The potentials of cochrane reviewers' comments and citation contexts in the recognition of randomized controlled trials' texts and their main sections. *Health Information Management*. 2020;17(4):181-8. [In Persian]
39. Times Higher Education. World University Rankings [Internet]. 2018 [cited 2018 Sep 26]. Available from: <https://www.timeshighereducation.com/world-university-rankings>
40. Sarker A, Molla D, Paris C. Automatic evidence quality prediction to support evidence-based decision making. *Artif Intell Med*. 2015;64(2):89-103.
41. Achananuparp P, Hu X, Shen X. The evaluation of sentence similarity measures. In: Song I-Y, Eder J, Nguyen TM, editors. *Data Warehousing and Knowledge Discovery. DaWaK 2008. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer; 2008. p. 305-16. (LNISA, volume 5182)
42. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc*. 1994;1(6):447-58.
43. Ruthven I, Lalmas M. A survey on the use of relevance feedback for information access systems. *Knowl Eng Rev*. 2003;18(2):1-55.
44. Cool C, Belkin N, Frieder O, Kantor P. Characteristics of text affecting relevance judgments. National online meeting. 1993;14:77-77.
45. Dhammi IK, Kumar S. Medical subject headings (MeSH) terms. *Indian J Orthop*. 2014;48(5):443-4.