Review Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2023 (4 Sep);37.95. https://doi.org/10.47176/mjiri.37.95

A Scoping Review of Adopted Information Extraction Methods for RCTs

Azadeh Aletaha[1,2], Leila Nemati-Anaraki[1,3]*, AbbasAli Keshtkar[4], Shahram Sedghi[1,5], Abdalsamad Keramatfar[6], Anna Korolyova[7,8,9]

Received: 25 Dec 2022 Published: 4 Sep 2023

Abstract

Background: Randomized controlled trials (RCTs) provide the strongest evidence for therapeutic interventions and their effects on groups of subjects. However, the large amount of unstructured information in these trials makes it challenging and time-consuming to make decisions and identify important concepts and valid evidence. This study aims to explore methods for automating or semi-automating information extraction from reports of RCT studies.

Methods: We conducted a systematic search of PubMed, ACM Digital Library, and Web of Science to identify relevant articles published between January 1, 2010, and 2022. We focused on published Natural Language Processing (NLP), machine learning, and deep learning methods that automate or semi-automate key elements of information extraction in the context of RCTs.

Results: A total of 26 publications were included, which discussed the automatic extraction of key characteristics of RCTs using various PICO frameworks (PIBOSO and PECODR). Among these publications, 14 (53.8%) extracted key characteristics based on PICO, PIBOSO, and PECODR, while 12 (46.1%) discussed information extraction methods in RCT studies. Common approaches mentioned included word/phrase matching, machine learning algorithms such as binary classification using the Naïve Bayes algorithm and powerful BERT network for feature extraction, support vector machine for data classification, conditional random field, non-machine-dependent automation, and machine learning or deep learning approaches.

Conclusion: The lack of publicly available software and limited access to existing software makes it difficult to determine the most powerful information extraction system. However, deep learning models like Transformers and BERT language models have shown better performance in natural language processing.

Keywords: Information extraction, NLP, Randomized Controlled Trials, automation

Conflicts of Interest: *None declared*
Funding: *This manuscript was developed as a part of Azadeh Aletaha PhD thesis which is monitored and funded by Iran University of Medical Sciences, Tehran, Iran with reference code:IR.IUMS.REC.1399.043.*

Cite this article as: Aletaha A, Nemati-Anaraki L, Keshtkar AA, Sedghi S, Keramatfar A, Korolyova A. A Scoping Review of Adopted Information Extraction Methods for RCTs. *Med J Islam Repub Iran.* 2023 (4 Sep);37:95. https://doi.org/10.47176/mjiri.37.95

Introduction

Evidence-Based Medicine (EBM) aims to teach individuals how to effectively utilize information and make informed decisions, even in the face of a large volume of available information. It emphasizes the integration of clinicians' experiences, patients' values, and the best scientific information that is currently available. The ultimate goal of

Corresponding author: *Dr Leila Nemati-Anaraki, Nematianaraki.l@iums.ac.ir*

1. Department of Medical Library and Information Science, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran
2. Evidence-Based Medicine Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran
3. Health Management and Economics Research Center, Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran
4. Department of Health Science Educational Development, School of Public Health, Tehran University of Medical Sciences. Tehran, Iran
5. Economics Research Center, Iran University of Medical Sciences, PO Box 14665-354, Tehran, Iran
6. Department of Data Analytics, Scientific Information Database (SID), Tehran, Iran
7. Computer Science Laboratory for Mechanics and Engineering Sciences (LIMSI), CNRS, Universit´e Paris-Saclay, F-91405 Orsay, France
8. School of Life Sciences and Facility Management Zurich University of Applied Sciences (ZHAW)
9. Fraser House, White Cross Business Park, Lancaster, LA1 4XQ

↑What is "already known" in this topic:
Information extraction systems rely on natural language processing and linguistic models, which play a crucial role in the information extraction process. NLP methods, tools, and more recently, deep learning transformers have been applied to automate or semi-automate the information extraction process in RCTs.

→What this article adds:
This article introduces the application of NLP, machine learning, and deep learning methods and tools to demonstrate automated or semi-automated methods in the information extraction process of RCTs. It highlights that Support Vector Machines (SVM) are more popular compared to other techniques. Additionally, it mentions that Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) have received more attention in information extraction within the context of Evidence-Based Medicine.

EBM is to enhance decision-making in clinical practice by ensuring that it is based on the most reliable and relevant evidence (1, 2).

Evidence-Based Medicine (EBM) utilizes a pyramid structure to classify different types of clinical evidence and assign them grades based on their strength. At the top of this pyramid are systematic reviews and meta-analyses, which are considered the highest level of evidence. These studies involve the comprehensive analysis of multiple randomized controlled trials (RCTs) to provide a more robust and reliable assessment of therapeutic interventions and their effects on groups of subjects. RCTs themselves are considered one of the strongest forms of evidence in EBM. By categorizing and grading different types of evidence, EBM helps clinicians make informed decisions based on the most reliable and rigorous research available (3, 4).

In the last decade, there has been a significant increase in the number of generated randomized controlled trials (RCTs) and systematic reviews of these trials (5). Systematic reviews are conducted to comprehensively review, evaluate, and synthesize all medical evidence pertaining to a specific research question and healthcare intervention (6, 7). To effectively process and analyze unstructured data, various information extraction approaches have been developed. These approaches aim to structure and extract valuable information from unstructured data. Natural language processing and linguistic models play a crucial role in the information extraction process. In the era of big data, we face numerous challenges due to the vast amount of data and its diverse structure. These challenges include managing and analyzing large volumes of data, dealing with data heterogeneity, and extracting meaningful insights from unstructured data. Information extraction techniques help address these challenges by enabling the extraction and organization of useful information from unstructured data, facilitating further analysis and decision-making.

According to the report of the International Data Corporation, it is projected that by 2020, unstructured data will make up 95% of global data. This indicates a significant increase in the volume and proportion of unstructured data compared to other types of data. The compound annual growth rate (CAGR) for unstructured data is estimated to be 65%, highlighting the exponential growth and importance of managing and extracting insights from unstructured data. This growth trend emphasizes the need for effective information extraction techniques and tools to process and analyze this vast amount of unstructured data (8, 9).

Randomized controlled trial texts contain valuable information for clinical research; In general, a significant part of the essential information of clinical trials is documented and stored with a large number of unstructured texts, making it difficult to effectively and accurately extract useful information. Furthermore, it can be time-consuming and costly to convert such unstructured texts into structured ones (10).

In recent years, natural language processing (NLP) and machine learning methods have been applied to automate the process of information extraction among the huge volume of texts and to facilitate the indexing of medical literature (11, 12).

Indeed, the filtering of trials and extracting relevant and precise information related to research questions and PICO elements can be a time-consuming and labor-intensive task. It often involves manually reviewing a large number of articles and extracting key information from them (13).

Absolutely, PICO elements and their related frameworks are indeed valuable for formulating search queries, particularly when searching for randomized controlled trials (RCTs) in clinical practice. PICO stands for Population, Intervention, Comparison, and Outcome. These elements provide a structured approach to formulating research questions and designing search queries that are specific and relevant to the clinical practice context. By clearly defining each element, researchers can narrow down their search and focus on finding RCTs that address their specific research question (14).

By leveraging these automated methods, researchers can save time and effort in filtering trials and extracting relevant information. This allows them to focus more on analyzing the extracted data and synthesizing the findings, leading to more efficient and reliable research outcomes.

While the automation of these tasks is still an ongoing area of research, the advancements in NLP and machine learning offer promising opportunities to alleviate the tediousness associated with trial filtering and information extraction, ultimately improving the efficiency of evidence synthesis and decision-making processes.

Reporting and extracting outcomes, especially the primary outcome, explains how the trial sample size was calculated (15-17). The process of information extraction is often time-consuming when researchers manually find key characteristics from articles to design an RCT protocol. As far as we are faced with structured and unstructured information in biomedical text, and this is challenging in practice to extract purpose-driven information to address specific clinical research questions and review research evidence to conduct research in clinical trials (18, 19). The methods of extracting information from biomedical texts are increasing which have been applied in the clinical field and specifically in systematic reviews of randomized controlled trials. One of the prominent technological approaches to information extraction is Natural Language Processing (NLP), including text mining and data extraction from different written resources (20).

Early attempts have been made for automatic knowledge extraction and mining from biomedical literature, and since the production of unstructured clinical trial data is fast and large-scale, it is extremely necessary to extract such textual data and generate further structured representations through automated approaches by applying NLP techniques (21-23).

Later, more advanced approaches using several natural language processing (NLP) techniques were used to automate the extraction of key features from randomized controlled trials. It could significantly reduce the time required for the design, conduct, and reporting of RCTs, thereby shortening the time it takes for evidence to be translated into clinical practice (24, 25).

[DOI: 10.47176/mjiri.37.95]

To open a new horizon for researchers in the field of randomized controlled trials, in this review, we investigated the NLP, Machine learning, and deep learning methods applied to demonstrate automating or semi-automating in the information extraction process in RCTs. In Evidence-Based Medicine, practitioners must access the best, relevant, and valid evidence in medical research, such as randomized controlled trials and systematic review and meta-analysis. So, the structure of these studies follows the PICO scheme (26, 27).

A significant outcome of this research has been the PICO (Population / Problem–Intervention–Comparison–Outcome) structure and its refined versions of PIBOSO, and PECODR frameworks to conduct research and design RCT protocols.

Research on how to automate the extraction of key features in randomized controlled trials (e.g., outcomes, ROB, or other key features) and software in use are limited. To fill the research gap, we identified existing methods related to the automated extraction of key elements in randomized controlled trials in biomedical texts for future works.

Methods

We used the PRISMA Extension (28) for methodological study (29). It was registered with registration DOI (10.17605/OSF.IO/2EZ5D) in the Open Science Framework (osf.io). This scoping review was guided by Arksey and O'Malley (30) and adopted by the 2017 Joanna Briggs Institute guidelines (31).

We performed a literature search using PubMed, ACM Digital Library, and Web of Science databases. The reason for choosing these databases is their high comprehensiveness in specialized computer science and biomedical issues. To collect the related documents in the field of information extraction, we use the following search query. Additionally, we reviewed the cited references of the included papers for further papers that matched our criteria.

TI= ((("Data Mining" OR (data AND mining) OR (text AND mining) OR (dataOR literature OR text) OR (mine? OR mining)) OR text mining-based OR (datamin* OR textmin*)) OR ("identification" OR "extraction" OR "extracting" OR "data extraction" OR detection OR "summarization" OR "learning approach" OR "automatically" OR Automatic OR automatically OR automation* OR summarization OR data OR information OR Keyword* OR text) OR ("Machine Learning" OR deep learning OR "supervised machine learning" OR "unsupervised machine learning" OR Transfer OR machine OR "learning algorithm*" OR "Interpreting" OR "Inferring" OR "classification"OR "Natural language processing" OR NLP OR question answering OR reading comprehension OR (term recognition or regular expression or regex))) AND TI= (BERT OR "Bidirectional Encoder Representations Transformer" OR BIOBERT OR SCIBERT OR ALBERT OR DistilBERT OR SpanBERT OR RoBERTa OR XLNet OR Transformer-XL) AND TI= ("medical evidence" OR "PICO" OR "PECODR" OR "intervention arms" OR "evidence synthesis" OR "experimental methods" OR "study design parameters" OR "Patient oriented Evidence" OR "eligibility criteria" OR Outcome extraction OR "clinical trial characteristics" OR "evidence based medicine" OR EBM OR "evidence based practice" OR "clinical trials" OR RCT OR "Randomized controlled trials" OR "Biomedical text" OR "Biomedical Evidence Synthesis" OR "clinical trial characteristics" OR clinical trial reports OR "clinical practice guidelines" OR living review).

Methodological analysis

Step 1. Identification of Research Question

The objective of this review is to find different types of methods used to extract the key features of RCT articles based on different types of PICO frameworks (PIBOSO and PECODR. Research on how to automate the extraction of key features in randomized controlled trials (e.g., outcomes, ROB, or other key features) and software in use are limited.

Research Questions

The research questions fall into two categories:

1-What types of methods and approaches are used to automate the extraction of key components from randomized controlled trials?

2-Which components have been automatically extracted based on the PICO, PIBOSO, and PECODR frameworks?

Eligibility criteria

To systematically review the literature on NLP, Machine learning, and deep learning approaches of randomized controlled trials, we defined these Eligibility criteria (inclusion and exclusion) as well as the search strategy and keywords (Table 1).

Table 1. Inclusion and exclusion criteria

Inclusion Criteria	
-The methods or results section recognized different frameworks such as outcome elements, PICO, PIBOSO, and PECODR structure from Randomized Controlled Trial Literature - Evaluate the accuracy, precision, recall, sensitivity, specificity, and/or F-measure, methods, algorithms, or tools that extract or label meta-information of text elements that may help in the extraction of information from these elements.	-full-text publications that describe an original NLP, Machine learning, and deep learning approach for extracting information related to randomized controlled trials - At least one entity was automatically extracted with evaluation results presented for that entity
Exclusion Criteria	
- The methods were not used for extracting data without the NLP, Machine learning, and deep learning approach for RCT - The report was an editorial, commentary, or another non-original research article. - The reports which have no evaluation components	

Step 2. Identifying relevant studies

Information Sources and searches

With the help of a medical librarian and an information specialist, search strategies were developed, and three databases were searched including PubMed, ACM digital library, and the Web of Science Core Collection. To broaden the scope of the search, Google Scholar was also used as a source for gray literature to find similar items.

Our searches were limited to the years 2010 to 2022. The reason for choosing this period was the emergence of the use of new automatic information extraction systems. The first group of keywords was related to information extraction methods. The second group of keywords was related to evidence synthesis and evidence-based medicine. The third group of keywords related to randomized controlled trials. All synonyms of keywords were checked in Medical Subject Heading (Mesh) available in the PubMed database.

The details of the search query and keywords are given in Appendix 1.

In total, we retrieved articles dealing with the labeled data. Table 2 illustrates information extraction methods, including 1) details of the algorithm class along with the extraction granularity used, the extraction source, dataset, and status of the project (Availability); 2) the core machine learning algorithms and the choice of feature extraction to use as input to the algorithm. Therefore, free access to the dataset allows researchers to use existing models in their work and to evaluate the results of their work in comparison to others 'studies. After searching the databases, the extracted articles were imported into Endnote Version X8.0.1 to organize, curate, and review their full text of them.

Step 3. Paper selections

Screening and selection of publications

We first removed the duplicates of the retrieved citations from the three resources based on the inclusion and exclusion criteria. The papers were checked by two authors independently of the aforementioned criteria. The included reports are classified into various categories according to the data elements attempted to be extracted from the original scientific articles. After checking all papers, the results were compared, and a Cohen's κ score for the inter-rater agreement was calculated. We resolved any disagreements between the two reviewers through discussion with the third author.

Step 4. Study selection

In total, 9331 articles were identified from three databases and reference lists of selected studies. After screening the studies based on their titles and abstracts, 9214 articles were irrelevant, and only 117 articles were selected for a more detailed review of abstracts. Of these articles, 12 were duplicate studies. Finally,26 articles met the inclusion criteria (Figure 1). The agreement on screening the abstracts and full texts was 0.97. The risk of bias assessment was not performed due to the type of review, which was scoping review (32).

Step 5. Charting the data

Two authors independently reviewed the full texts of 26

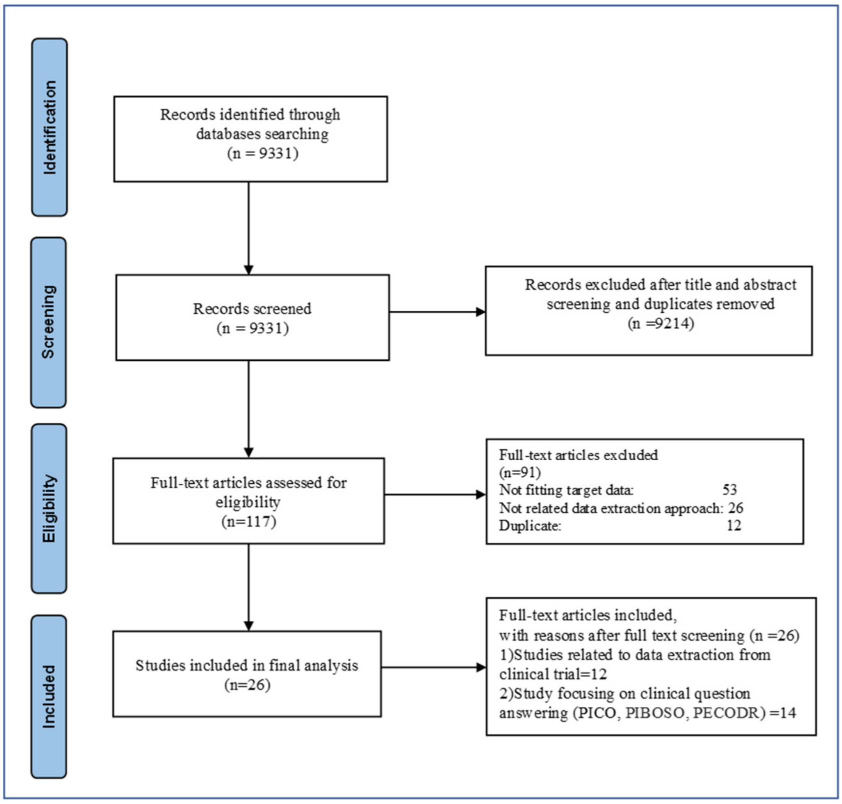

Figure 1. PRISMA flowchart

[DOI: 10.47176/mjiri.37.95]

articles to extract data, including the particular entity automatically extracted by the study, algorithm or technique employed, and evaluation of results, into a data abstraction spreadsheet. We resolved any disagreements through consensus with the third author. PICO, PIBOSO, and PECODR frameworks were considered to obtain data elements. several characteristics were recorded for each theme listed below:

Publication, Methods, Size/type/source, Classes, Availability, Assessment, and limitation.

Results

To gain insight into the kind and extent of work done in the field of NLP in randomized controlled trials, we extracted the following information from the papers: Software used; classes; NLP methods; dataset; availability, and performance measures of the reported data extraction Method. Table 2 presents a list of items and the key features of the selected articles in the information extraction process based on PICO, PIBOSO (Population–Intervention–Background–Outcome–Study Design–Other), and PECODR ((clini-cal) Patient, Exposure, Comparison, Outcome, Duration, Results) frameworks. This provides the types of methods, extraction level, and new approaches applied to extract key features and published methods to extract. For the main NLP methods used in the reviewed papers, we rec-

Table 2. A summary of included information extraction methods

Publication	Methods	Size/type/source	Classes	Availability	Assessment	limitation
(33)	SVM [1]/ MLP [2] /RF [3]/ NB [4] supervised classification algorithms, Auto-labelled structured abstracts, sentence level	26,000 Abstracts, Medline/ PubMed	PICO(I/C)	-	10-fold cross-validation. F-score P: 86.3%, I/C: 67% O:56.6%	The task complexity, use of non-PICO-specific vocabulary, and sentence heading outcome refer in more than one sentence. The O or I elements are more difficult to identify than P elements.
(34)	Robust statistical classification approach in two levels of classification (identify each PICO element in the document,2-make a coarser-grain annotation to annotate a sentence as describing only one of the PICO elements	151,646 Abstracts/PubMed	(P, IC, O)	-	10-fold cross-validation. F1- P: 77.8% I:68,3% O: 50%	-
(35)	Naïve Bayesian (NV)	23,472 Abstracts/PubMed	(P-I-O)	-	Ten-fold cross-validation F-P: 0.91% I:0.75% O: 0.88%	-
(36)	CRF [5]	1,000 Abstracts/Medline	(PIBOSO)	https://github.com/olab-knbit/ebm-sentence-classification	F-scores P:80.9% I:66.9 %, O:63.1 %	-
(37)	NLTK, NB classifier	19,854 Abstracts/Medline	(P-I-O)	-	Ten-fold cross-validation F-score P:73.9%/ I:66.2%/ O:73.1	no manual review in answering EBM questions with PICO.
(38)	Generic rule-based approach	60 + 30 Abstracts/Medline	(P, O, Exposure, covariates, and Effect size)	http://gnteam.cs.manchester.ac.uk/ old/epidemiology/home.html	F1 score: 93.3 for P 82.4 for O	1-The current work does not include the identification of synonymous expressions or more detailed mapping of identified terms to existing knowledge repositories. 2-focused only on abstracts rather than full-text articles.

[1] Support Vector Machines (SVM)
[2] Multi- Layer Perceptron (MLP)
[3] Random Forests (RF)
[4] Naive Bayes (NB)
[5] Conditional random fields (CRF)

Table 2. Continued

Publication	Methods	Size/type/source	Classes	Availability	Assessment	limitation
(39)	Hybrid approach (MLMs (CRF) and RBMs) used in cTAKES[6]	3000 abstracts/ PubMed	(PICO)	-	-	-
(13)	Labeled via supervised distant supervision[7]	12808 full texts per class), 50 + 133 manually annotated for evaluation / CDSR[8]	(PICO)	-	cross-fold validation pairwise κ = 0.74, overall, and κ = 0.81 per-article AUC, P:94.7 I: 93.6 O:90	-
(40)	In sentence ranking (ML model) and NLP approach. In fragment-level extraction (regular expression matching, mapping to UMLS concepts, and element-specific dictionary)	48 full texts in 8 systematic reviews/Cochrane library	(Sample size, group size, PICO	-	F1 score: for Sample Size/Group size:90.3/ P:79.8 Study arm:86.8/O:81.8	This study focused on sample size and PICO elements, which are commonly reported in RCT studies. other machine learning models, such as linear regression, multilayer perceptron, and Gaussian processes that were not evaluated in this study
(41)	A (CRF) and (LSTM) neural tagging model	5000 Abstracts/ Medline/ PubMed	(PICO)	https://www.ccs.neu.edu/home/bennye/EBM-NLP/pubs.html /	CRF :F1 score:P:0.5II:0.32 O:0.29 LSTM-CRF: F1 score P:0.71 I:0.65 O:0.63	Detailed but small (hundreds of documents) and large but distant (paragraph-level labels)
(42)	LSTM-based ANN [9] architecture	489,026 Abstracts/PubMed/Medline	(PICOM)	https://github.com/jind11/LSTM-PICO-Detection	F1 score: P:85.6 I:78.1 O:83.8 M:85.6	-
(43)	LSTM-CRF model	170 abstracts/PubMed/Medline	(PICO)	https://github.com/Tian312/PICO_Parser	F1 score P:0.75 I:0.61. O:0.56	-
(44)	(RNNs)/ BiLSTMs BERT). 1-PICO Entity Recognizer (Recursive Neural Networks (RNNs) for character feature extraction and 2-PICO sentence classifier	5000 abstracts PubMed/Medline	(PICO)	https://github.com/nstylia/pico_entities/	10-fold cross-validation. F1score for P:80 I:65 O:78	-
(45)	sentence annotations without any span annotations BLUE and BERT neural language models	500 abstracts/PubMed	(PICO)	https://github.com/evidence-surveillance/sent2span	F1 score: P:0.84 I:0.83 O:0.83	-

[6] Clinical Text Analysis and Knowledge Extraction System(cTAKES)
[7] supervised distant supervision (SDS)
[8] Cochrane Database of Systematic Reviews (CDSR)
[9] Artificial Neural Network (ANN)

orded the performance expressed values of recall, precision, and F-measure.

Table 3 provides a summary of the information extraction methods used to extract key features from randomized controlled trials (RCTs). This information can be valuable for researchers, especially when conducting a systematic review of RCTs and assessing the risk of bias in the included articles. By automating the process of extracting key characteristics from each RCT, particularly when it comes to outcome extraction, researchers can efficiently identify different patient outcome reports and assess outcome diversity. This can be particularly useful in designing RCT protocols, as it helps researchers understand the range of outcomes reported in similar studies and incorporate a comprehensive set of outcomes in their own protocol. By using automation and rigorous extraction methods, researchers can save time and effort in manually extracting and analyzing key features from RCTs, allowing for a more efficient

[DOI: 10.47176/mjiri.37.95]

Table 3. A summary of included information extraction methods in Randomized Controlled Trials

Publication	Method	Class/Type/Size	Evaluation	Availability
(48)	Machine learning Heuristic. An information extraction (IE) engine searches articles for text fragments. (Uses a statistical text classifier, (SVM)[1], (HMM)[2], and (CRF)[3],	(PICO)/RCTs/ 21RCTs abstracts and full texts: a set of 1050 tasks in132+50 articles from 25 journals	Precision and recall Eligibility criteria:1.00 Sample size:0.89, 0.87 Primary outcome name:0.97 Primary outcome time point: 0.90, Secondary outcome name: 0.93 Duration of treatment:0.84	-
(49)	Machine learning (heuristic features). CRF classifier and MALLET Simple Tagger	Treatment Group, Outcome. 263 RCTs of the British Medical Journal (BMJ)	F1 score Treatment group:0.76 Outcomes:0.42	-
(50)	Using an automated Sequence Annotation Pipeline provides an interface for querying biomedical knowledge sources and integrating the results plans	Statistical analysis. (Outcome measure). 42 full-text RCTs related to chemotherapy of non-small cell lung cancer PubMed Central	precision, recall, and F-score (introduction: 0.86), (outcomes, sample size: 0.84)	-
(51)	The core of the system is completely based on statistical techniques. consists of two components: a basic classifier and an inference procedure. A Maximum Entropy classifier is first trained by using a standard set of linguistic features.	PICO/99 RCT Abstracts	F1 Score P:0.88 I:0.72, Control arms 0.64. O: 0.72. The overall precision of the system is 0.68	https://github.com/antoniotre86/IERCT
(52)	Rule-based approach, SVMs	Sample Size/200 RCTs Abstracts	Using 10-fold -cross-validation. The best accuracy score obtained on the training set is 94%.	-
(53)	A novel variant of Convolutional Neural Networks (CNNs) was adapted for text classification. Using several machines learning (ML) data-extraction models	PICO/ROB[4]/RCTs Fulltext	The accuracy of the overall classification of articles as describing high/unclear or low-risk RCTs achieved by our model remained 5–10 points lower than that achieved in published (human-authored)	https://github.com/ijmarshall/robot reviewer
(54)	(1) clinical entity and attribute recognition, (2) negation detection, (3) relation extraction, and (4) concept normalization and output structuring.	230 Alzheimer's RCTs/ Eligibility criteria	In task-specific evaluations, the best F1 score for entity recognition was 0.79, and for relation, extraction was 0.89	https://github.com/Tian312/EliIE
(55)	Deep learning models (BERT, SciBERT, BioBERT) . rules based on syntactic structure provided by spaCy dependency parser, a combination of bi-LSTM, CNN, and CRF using GloVe, word embeddings, and character-level	outcome extraction, significant level, and relation extraction	F1 score O:79.42 Relation extraction:94 Significance levels: 97.86%	https://zenodo.org/record/3234834

[1] Support Vector Machines (SVM)
[2] Hidden Markov Models (HMM)
[3] Conditional Random Fields (CRF)
[4] Risk of Bias Assessment

and thorough systematic review process. This can ultimately enhance the quality and reliability of evidence-based research in the field of clinical trials.

A description of clinical study design is often used to classify the types of evidence generated (46). In the early design of a randomized controlled trial, identifying and extracting the appropriate outcomes and other key features of a trial can potentially aid in determining the sample size to conduct a randomized control trial (47). However, in this study, we did not intend to present the best method and approach but to have an overview of information extraction methods such as machine learning, deep learning, and natural language processing (NLP) techniques. The field of NLP is growing rapidly to become one of the most active

[DOI: 10.47176/mjiri.37.95]

Table 3. Continued

Publication	Method	Class/Type/Size	Evaluation	Availability
(56)	Machine learning and rule-based methods to extract information from the RCT abstracts and PICO elements and map these snippets to normalized Mesh vocabulary terms.	Mesh labels and PICO concepts, Risk of bias, Sample size/304 111 RCTs registrations from the International Clinical Trials Registry Platform and World Health Organization International Clinical Trials Registry Platform	F1 scores P:0.71 I:0.65 O: 0.63.	https://trialstreamer.ieai.robotreviewer.net/
(57)	NLP techniques/ The evidence extraction pipeline is composed of four primary phases. First, text snippets that convey information about the trial's treatments (or interventions), outcome measures, and results are extracted from abstracts. Finally, the clinical concepts expressed in the extracted spans are normalized to a structured vocabulary to ground them in an existing knowledge base and allow for aggregations across the trial.	ICO/RCTs	Macro-averaged scores for ICO span prediction. F1 Score:0.67	https://github.com/bepnye/evidence_extraction/
(58)	Rule-based methods and Machine learning methods (deep learning) for similarity statements and within-group comparisons). The language representations that were tested include: BERT BioBERT and SciBERT trained on the BERT corpus and a scientific corpus of 3.1B words	Reported outcomes and statistical significance levels/180 RCTs abstract	F1 score Primary outcome:88.4 Reported outcome:79.4 Outcome similarity assessment: 89.75/ Similarity statements extraction:82.4 significance levels:97.86	https://github.com/aakorolyova/
(59)	Spans describing interventions and snippets that report key results. In a second step, link the identified evidence-bearing snippet to the extracted outcome and intervention to which it most likely pertains. Extract, Link, infer (ELI) approach. A linear classification layer is fine-tuned on top of SciBERT that predicts the directionality of the finding O/I	ICO/RCTs full text	F1 score O:0.78% I:0.75% C:0.70%	-

research areas in trial studies that extract key features, including different PICO elements and especially outcomes, which is the basis for determining the sample size for RCT design (43).

Therefore, a more practical approach and techniques for analyzing data related to automated information extraction of key characteristics of trials and for evidence synthesis are required to improve the RCT protocol design.

Figure 2 shows the distribution use of main NLP and machine learning methods in information extraction over the papers, as well as associated methods and techniques used, is shown in Figure 2. The main method that has been applied most is rule-based, including Regex methods (n=10). It demonstrates the system architectures implemented in the included publications. An architecture combining a word embedding + long short-term memory (LSTM) network would have been divided into the two sub-components. The binary classifiers were grouped into two-components naive Bayesian and bidirectional encoder representation decision trees (BERT). Since SVM is also a binary classifier, it was assigned as a separate category due to its popularity. The final classifications are a mixture of non-machine-leaning automation (application programming interface (API) and metadata retrieval, PDF extraction, rule base), machine-learning (naïve Bayes, decision trees, SVM), and neural or deep learning approaches (convolutional neural network, LSTM, transformers, or word embeddings). However, there is no consensus pointing out the use of these architectures in the design of automatic information extraction systems.

[DOI: 10.47176/mjiri.37.95]

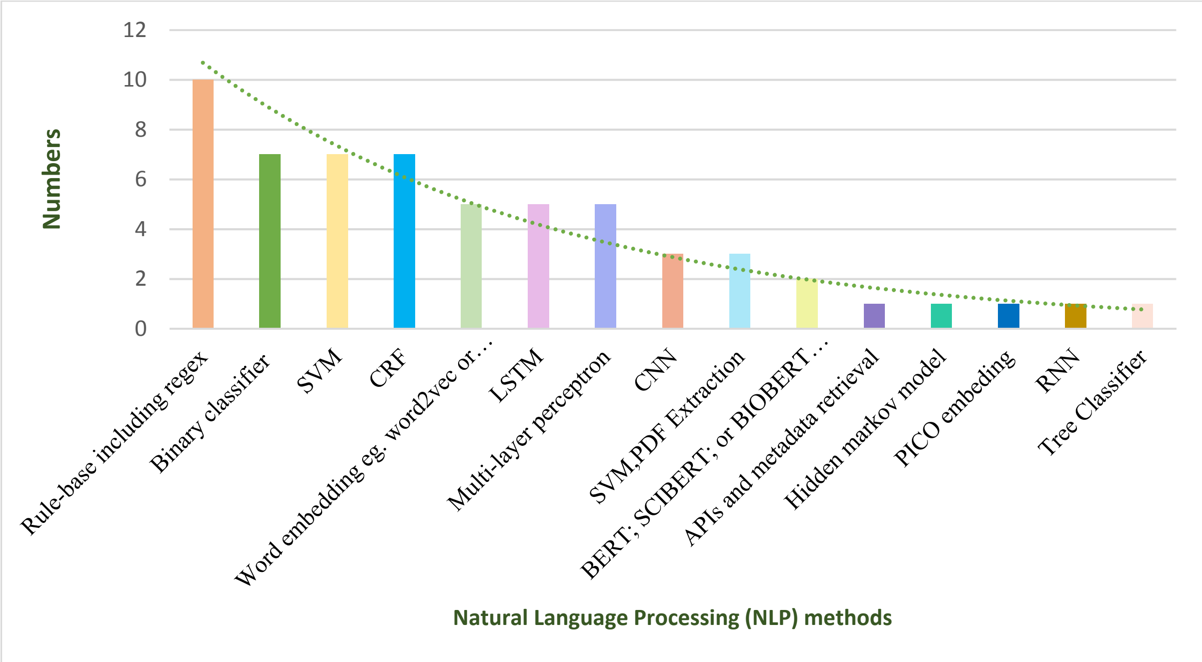

Figure 2. Methods used for automating information extraction in the included publications

Binary classifiers, specifically Naïve Bayes and SVM, are the most commonly used system components for information extraction. These classifiers are currently used in most studies. Rule bases, including heuristic, word list, and regular expression approaches, were one of the first techniques used for data extraction in the EBM literature. It remains one of the most widespread automation approaches. Automation systems implement rule bases to identify phrases for entities, such as exposure, effect size, and covariate, and combine them with entity-level machine learning classifiers, such as patients, intervention, and outcome (primary or secondary) extracted from sentences. In recent years, embedding and neural architectures are increasingly used in automating. LSTM, CNN, and Recurrent neural networks (RNN) have received more attention in information extraction in Evidence-Based medicine.

Discussion

The purpose of the present study was to identify and describe the use of NLP and machine learning methods for information extraction in randomized controlled trials (RCTs) from 2010 to 2022. With the advancements in data science and natural language technologies, as well as the increasing automation of evidence synthesis and information extraction from structured and unstructured biomedical data, significant changes have occurred in the field of information retrieval and big data.

By leveraging NLP and machine learning techniques, researchers, particularly those involved in systematic reviews and RCTs, can benefit in several ways. Firstly, these methods can save time and costs by automating the process of extracting relevant information from RCTs. Manual extraction can be time-consuming and prone to errors, but with the use of NLP and machine learning, researchers can extract and analyze data more efficiently.

Additionally, these methods can improve data-driven decision-making processes. By extracting and synthesizing information from RCTs, researchers can gain valuable insights and make evidence-based decisions. This can enhance the quality and reliability of research findings, leading to better-informed healthcare interventions and policies.

Overall, the integration of NLP and machine learning methods in information extraction from RCTs has the potential to revolutionize the field of systematic reviews and evidence synthesis. It offers opportunities to save time, reduce errors, and improve decision-making processes, ultimately advancing the field of healthcare research.

Our review highlighted the new NLP and machine learning methods and approaches in information extraction from trials and in question and answering systems based on different frameworks, such as PICO elements. Rule-based approaches are most frequently used, and there is a trend toward using neural networks such as the bidirectional training of transformers and different BERT language models. Most of the publications, which were reviewed, focused on extracting information from abstracts.

A few articles extracted information from full texts of Randomized controlled trial studies (n=9, 34%), but the information extracted on this issue is still sparse, and little research has been done in this area. Fourteen studies explored the extraction of interventions and outcomes (13, 33-35, 37, 41, 48, 51, 53, 55-57, 59-62).

None of the studies used the same corpus. Only two studies extracted the essential data elements from outcome measures and divided them into primary and secondary outcomes. For example, Kiritchenko et al. were able to achieve an f-score of 0.97% for primary outcome data elements and 0.93% for secondary outcome data elements on a dataset of 50 full-text journal articles (48). Koroleva et al. achieved an f-score of 88.42 for primary outcome data elements on a dataset of 180 full-text journal articles and did not extract secondary outcomes (58). The availability of the final tools was very poor. We found that only 7.6% of all publications were based on available tools for their data extraction system and had a graphical user interface.

Previous reviews on the automation of data extraction in systematic review processes describe methods and new approaches. Schmidt et al. focus on the data extraction methods for different systematic reviews and evidence-based publications describing data extraction for interventional studies (25). Tsafnat et al. described the information systems for automation of each stage of systematic review (63). We concentrated on information extraction on randomized controlled trials and outcomes and PICO data elements. None of the existing reviews focus on the information extraction step to conduct an RCT and systematic review of trials (25, 63, 64). For example, Schmidt provided a broad overview of published methods and tools aimed to automate or semi-automate the data extraction process in the context of a systematic review of medical research studies (25).

In comparison, we identified 26 studies and classified and summarized current methods and tools in automation of critical characteristics of Randomized control trials and systematic review of trials due to the importance of clinical trial studies in recent decades and significant changes in their methodology (65)

We have provided added value for the new methods in extracting critical features of randomized controlled trials, especially the extraction of the reported outcome. Wallace et al. suggested an active learning framework for reducing the workload in citation screening for inclusion in the systematic reviews (66). Nye et al. introduced Trial Streamer, a living database of clinical trial reports to extract critical pieces of information from biomedical abstracts that clinicians need when conducting a risk of bias assessment of the literature. It also removes the description of participants in the trial, the treatments compared in each arm, and the outcomes measured. It attempts to infer which interventions were reported to work best by determining their relationship with identified trial outcome measures (57).

Koroleva et al propose a Natural Language Processing (NLP) system for detecting several types of spin in biomedical articles reporting randomized controlled trials (RCTs) and an aid tool for assisting both authors and peer reviewers in detecting potential spin. Overinterpretation of research results, also known as distorted reporting or spin, is a serious issue in research reporting (58).

There is no gold standard or dataset for evaluation. This makes it very difficult to claim which methods are more effective. And most of these methods focus on the risk of bias assessment of studies in conducting a systematic review and not merely extracting key characteristics of information to execute and design an RCT and or a systematic review of the trial. However, Due to the interdisciplinary nature and multiplicity of automatic information extraction and thematic dispersion from the systematic review and randomized controlled trial studies, it is not easy to present a clear path of the trends and approaches in this issue.

We believe that developing information extraction methods in conducting a systematic review of trials and RCT would provide valuable insights for scholars, clinicians, and other healthcare professionals in this field.

Conclusion

Our Methodological review describes the methods and measurements in information extraction automation of key characteristics of RCT and only a few studies that have reported their prototype system available. Information extraction is the task of automatically identifying important key characteristics in unstructured natural language text. It involves several subtasks, including named entity recognition, event extraction, and relation extraction (67).

In this survey, we reviewed recent studies that focus on the applications of information extraction techniques for the processing of randomized controlled trial data. We attempt to encourage researchers to seek the potential to combine advanced deep learning techniques and Methodology, including deep reinforcement learning, deep neural networks, BERT models, and convolutional neural networks, with NLP techniques to deal with issues regarding randomized controlled trials (68).

Deep learning models such as bidirectional encoder representations from Transformers are getting popular reported in recent studies and

their major building block is transformers to learn contextual relations between words in sentences (22, 69, 70).

This makes it very difficult to draw conclusions on which is the best-performing system. Many of them were not available, and few publications made their datasets available to the public. Some datasets and codes were available on GitHub , and their prototypes were evaluated. And also, information extraction is a complicated task and requires Subject-matter experts. However, we hope these automated extraction methods aid researchers in designing an RCT protocol and help them in the risk of bias assessment of systematic review of trials.

This study also provides a deeper insight into information extraction research. Our analysis shows that there has been significant growth in this field until 2022.NLP, machine learning, deep learning, and BERT-based Embeddings are to be the next frontier topics in this area.

Limitations

There are some limitations in this study. The WOS core database has collected only some of the newly added research articles that are cited daily in WOS. But research demonstrates that there is a high overlap between WOS and Scopus databases for analysis in computer science and natural science (71). There is the likelihood that information extraction algorithms and evidence synthesis tools were not published in the journals we searched, or we might have missed some of them.

Credit authorship contribution statement

Conceptualization designed the analysis, methodology and performed the analysis and Writing – original draft: Azadeh Aletaha

Supervision and writing – review & editing: Leila Nemati-Anaraki
Formal analysis: Abbas Ali Keshtkar
Project administration: Abdal Samad Keramatfar
Visualization and Data Curation: Shahram Sedghi
Validation: Anna Korolyova

Conflict of Interests
The authors declare that they have no competing interests.

References
1. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. BMJ. 1995;310(6987):1122-6.
2. Williams H, Bigby M, Diepgen T, Herxheimer A, Naldi L, Rzany B. Evidence-based dermatology: John Wiley & Sons; 2009.
3. Rosner AL. Evidence-based medicine: revisiting the pyramid of priorities. J Bodyw Mov Ther. 2012;16(1):42-9.
4. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21(4):125-7.
5. Fontelo P, Liu F. A review of recent publication trends from top publishing countries. Syst Rev. 2018;7(1):147.
6. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev. 2019;10:Ed000142.
7. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf MI, et al. Searching for and selecting studies. Cochrane Handbook for systematic reviews of interventions. 2019:67-107.
8. Gantz J, Reinsel D. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC iView: IDC Analyze the future. 2012;2007(2012):1-16.
9. Adnan K, Akbar R. Limitations of information extraction methods and techniques for heterogeneous unstructured big data. IJEBM.2019;11:1847979019890771.
10. Jung H, Hong S, Park J, Park M, Sun J, Lee S, et al. MA19. 06 Successful Development of Realtime Automatically Updated Data Warehouse in Health Care (ROOT-S). JTO. 2019;14(10):S328.
11. Schmidt L, Weeds J, Higgins J. Data mining in clinical trial text: Transformers for classification and question answering tasks. ArXiv:200111268. 2020.
12. Huang M, Névéol A, Lu Z. Recommending MeSH terms for annotating biomedical articles. JAMIA. 2011;18(5):660-7.
13. Wallace BC, Kuiper J, Sharma A, Zhu MB, Marshall IJ. Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision.. JMLR. 2016;17.
14. Brockmeier AJ, Ju M, Przybyła P, Ananiadou S. Improving reference prioritisation with PICO recognition. BMC Med Inform Decis Mak. 2019;19(1):256.
15. Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. Lancet. 2005;365(9465):1159-62.
16. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. BMJ. 2017;357:j2490.
17. Hopewell S, Dutton S, Yu LM, Chan AW, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. BMJ. 2010;340:c723.
18. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7(9):e1000326.
19. Luo L, Li L, Hu J, Wang X, Hou B, Zhang T, et al. A hybrid solution for extracting structured medical information from unstructured data in medical records via a double-reading/entry system. BMC Med Inform Decis Mak. 2016;16(1):114.
20. Hearst MA, editor Untangling text data mining. Proceedings of the 37th Annual meeting of the Association for Computational Linguistics; 1999.
21. Schmidt L, Olorisade BK, McGuinness LA, Thomas J, Higgins JPT. Data extraction methods for systematic review (semi)automation: A living review protocol. F1000Res. 2020;9:210.
22. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag. 2018;13(3):55-75.
23. Zhu F, Patumcharoenpol P, Zhang C, Yang Y, Chan J, Meechai A, et al. Biomedical text mining and its applications in cancer research. J Biomed Inform. 2013;46(2):200-11.
24. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Syst Rev. 2019;8(1):163.
25. Schmidt L, Olorisade BK, McGuinness LA, Thomas J, Higgins JPT. Data extraction methods for systematic review (semi)automation: A living systematic review. F1000Res. 2021;10:401.
26. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club. 1995;123(3):A12-3.
27. 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.
28. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
29. Mbuagbaw L, Lawson DO, Puljak L, Allison DB, Thabane L. A tutorial on methodological studies: the what, when, how and why. BMC Med Res Methodol. 2020;20(1):226.
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.
31. Peters M, Godfrey C, Khalil H, Mcinerney P, Soares C, Parker D. 2017 guidance for the conduct of JBI scoping reviews. Joana Briggs Inst Rev Man. 2017;13:141-6.
32. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169(7):467-73.
33. 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;10:29.
34. Boudin F, Shi L, Nie J-Y, editors. Improving medical information retrieval with pico element detection. Advances in Information Retrieval: 32nd European Conference on IR Research, ECIR 2010, Milton Keynes, UK, March 28-31, 2010 Proceedings 32; 2010: Springer.
35. Huang K-C, Liu CC-H, Yang S-S, Xiao F, Wong J-M, Liao C-C, et al., editors. Classification of PICO elements by text features systematically extracted from PubMed abstracts. 2011 IEEE International Conference on Granular Computing; 2011: IEEE.
36. Kim SN, Martinez D, Cavedon L, Yencken L. Automatic classification of sentences to support Evidence Based Medicine. BMC Bioinformatics. 2011;12 Suppl 2(Suppl 2):S5.
37. Huang KC, Chiang IJ, Xiao F, Liao CC, Liu CC, Wong JM. PICO element detection in medical text without metadata: are first sentences enough? J Biomed Inform. 2013;46(5):940-6.
38. Karystianis G, Buchan I, Nenadic G. Mining characteristics of epidemiological studies from Medline: a case study in obesity. J Biomed Semantics. 2014;5:22.
39. Chabou S, Iglewski M, editors. PICO Extraction by combining the robustness of machine-learning methods with the rule-based methods. 2015 World Congress on Information Technology and Computer Applications (WCITCA); 2015: IEEE.
40. Bui DDA, Del Fiol G, Hurdle JF, Jonnalagadda S. Extractive text summarization system to aid data extraction from full text in systematic review development. J Biomed Inform. 2016;64:265-72.
41. Nye B, Jessy Li J, Patel R, Yang Y, Marshall IJ, Nenkova A, et al. A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature. Proceedings of the conference Association for Computational Linguistics Meeting. 2018;2018:197-207.
42. Jin D, Szolovits P, editors. Pico element detection in medical text via long short-term memory neural networks. Proceedings of the BioNLP 2018 workshop; 2018.
43. Kang T, Zou S, Weng C. Pretraining to Recognize PICO Elements from Randomized Controlled Trial Literature. Stud Health Technol Inform. 2019;264:188-92.
44. Stylianou N, Razis G, Goulis DG, Vlahavas I. EBM+: Advancing Evidence-Based Medicine via two level automatic identification of Populations, Interventions, Outcomes in medical literature. Artif Intell Med. 2020;108:101949.

45. Liu S, Sun Y, Li B, Wang W, Bourgeois FT, Dunn AG. Sent2Span: span detection for PICO extraction in the biomedical text without span annotations. arXiv preprint arXiv:210902254. 2021.
46. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet. 2002;359(9300):57-61.
47. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med. 2007;33(2):155-61.
48. 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;10:56.
49. Summerscales RL, Argamon S, Bai S, Hupert J, Schwartz A, editors. Automatic summarization of results from clinical trials. 2011 IEEE International Conference on Bioinformatics and Biomedicine; 2011: IEEE.
50. Hsu W, Speier W, Taira RK. Automated extraction of reported statistical analyses: towards a logical representation of clinical trial literature. AMIA Annual Symposium proceedings AMIA Symposium. 2012;2012:350-9.
51. Trenta A, Hunter A, Riedel S. Extraction of evidence tables from abstracts of randomized clinical trials using a maximum entropy classifier and global constraints. arXiv preprint arXiv:150905209. 2015.
52. Sarker A. Automated Extraction of Number of Subjects in Randomised Controlled Trials. arXiv preprint arXiv:160607137. 2016.
53. Marshall IJ, Kuiper J, Banner E, Wallace BC. Automating Biomedical Evidence Synthesis: RobotReviewer. Proceedings of the conference Association for Computational Linguistics Meeting. 2017;2017:7-12.
54. Kang T, Zhang S, Tang Y, Hruby GW, Rusanov A, Elhadad N, et al. EliIE: An open-source information extraction system for clinical trial eligibility criteria. JAMIA. 2017;24(6):1062-71.
55. Koroleva A, Paroubek P, editors. Extracting relations between outcomes and significance levels in Randomized Controlled Trials (RCTs) publications. Proceedings of the 18th BioNLP Workshop and Shared Task; 2019.
56. Marshall IJ, Nye B, Kuiper J, Noel-Storr A, Marshall R, Maclean R, et al. Trialstreamer: A living, automatically updated database of clinical trial reports. JAMIA. 2020;27(12):1903-12.
57. Nye BE, Nenkova A, Marshall IJ, Wallace BC. Trialstreamer: Mapping and Browsing Medical Evidence in Real-Time. Proceedings of the conference Association for Computational Linguistics North American Chapter Meeting. 2020;2020:63-9.
58. Koroleva A, Kamath S, Bossuyt P, Paroubek P, editors. DeSpin: a prototype system for detecting spin in biomedical publications. Proceedings of the 19th SIGBioMed Workshop on Biomedical Language Processing; 2020.
59. Nye BE, DeYoung J, Lehman E, Nenkova A, Marshall IJ, Wallace BC. Understanding Clinical Trial Reports: Extracting Medical Entities and Their Relations. AMIA Joint Summits on Translational Science proceedings AMIA Joint Summits on Translational Science. 2021;2021:485-94.
60. Kang T, Zou S, Weng C. Pretraining to recognize PICO elements from randomized controlled trial literature. Studies in health technology and informatics. 2019;264:188.
61. Boudin F, Nie J-Y, Dawes M, editors. Clinical information retrieval using document and PICO structure. Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics; 2010.
62. Jin D, Szolovits P. Advancing PICO element detection in biomedical text via deep neural networks. Bioinformatics. 2020;36(12):3856-62.
63. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. Syst Rev. 2014;3:74.
64. Jonnalagadda SR, Goyal P, Huffman MD. Automating data extraction in systematic reviews: a systematic review. Syst Rev. 2015;4:78.
65. Sessler DI, Imrey PB. Clinical Research Methodology 3: Randomized Controlled Trials. Anesth Analg. 2015;121(4):1052-64.
66. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. BMC bioinformatics. 2010;11:55.
67. Klein D, Smarr J, Nguyen H, Manning CD, editors. Named entity recognition with character-level models. Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003; 2003.
68. Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. J Mach Learn Res. 2011;12(ARTICLE):2493− 537.
69. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.
70. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.
71. Mongeon P, Paul-Hus A. The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics. 2016;106:213-28.

[DOI: 10.47176/mjiri.37.95]