Natural language processing systems for diagnosing and determining level of lung cancer: A systematic review

Mahdieh Montazeri¹ , Ali Afraz¹ , Raheleh Mahboob Farimani², Fahimeh Ghasemian^{3*}

¹Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

²Department of Health Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran

³Department of Computer Engineering, Faculty of Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

Article Info

Article type:
Review

Article History:

Received: 2020-12-09

Accepted: 2021-03-08

Published: 2021-03-19

* Corresponding author:

Fahimeh Ghasemian

Department of Computer
Engineering, Faculty of Engineering,
Shahid Bahonar University of
Kerman, Kerman, Iran

Email:

ghasemianfahime@uk.ac.ir

Keywords:

Lung Cancer

Natural Language Processing

Diagnose

Staging

ABSTRACT

Introduction: Lung cancer is the second most common cancer for men and women. Using natural language processing to automatically extract information from text, lead to decrease labor of manual extraction from large volume of text material and save time. The aim of this study is to systematically review of studies which reviewed NLP methods in diagnosing and staging lung cancer.

Material and Methods: PubMed, Scopus, Web of science, Embase was searched for English language articles that reported diagnosing and staging methods in lung cancer Using NLP until DEC 2019. Two reviewers independently assessed original papers to determine eligibility for inclusion in the review.

Results: Of 231 studies, 7 studies were included. Three studies developed a NLP algorithm to scan radiology notes and determine the presence or absence of nodules to identify patients with incident lung nodules for treatment or follow-up. Two studies used NLP to transform the report text, including identification of UMLS terms and detection of negated findings to classifying reports, also one of them used an SVM-based text classification system for staging lung cancer patients. All studies reported various performance measures based on the difference between combinations of methods. Most of studies have reported sensitivity and specificity of the NLP algorithm for identifying the presence of lung nodules.

Conclusion: Evaluation of studies in diagnosing and staging methods in lung cancer using NLP shows there is a number of studies on diagnosing lung cancer but there are a few works on staging that. In some studies, combination of methods was considered and NLP isolated was not sufficient for capturing satisfying results. There are potentials to improve studies by adding other data sources, further refinement and subsequent validation.

Cite this paper as:

Montazeri M, Afraz A, Mahboob Farimani R, Ghasemian F. Natural Language Processing Systems for Diagnosing and Determining Level of Lung Cancer: A Systematic Review. Front Health Inform. 2021; 10: 68. DOI: [10.30699/fhi.v10i1.264](https://doi.org/10.30699/fhi.v10i1.264)

INTRODUCTION

Today cancer is one of the main health issues all over the world [1-4]. The lung cancer is one of the most common cancers worldwide and is the major reason for mortality from cancer in the world [5-9]. Lung cancer is assessed to present 228,150 new cases and 142,670 deaths in 2019 in the United States [10]. Late diagnosis and treatment in lung cancer can result from a failure to act upon abnormal radiologic findings in a timely way [11]. The most common cause of lung cancer is the lengthy exposure to tobacco smoke, which is the explanation of 90% of lung

cancers [12-14]. The Percent of lung cancer in individuals who do not smoke is 15% and the reason is because of a combination of factors including genetic factors, radon gas, asbestos, and air contamination such as cigarette smoke of another person [15]. Stage and treatment modality have been the main factors for lung cancer prognosis [16]. The stage of a cancer groups its progression, in terms of the size and location of the primary tumor, just as any spreading to lymph nodes or formation of distant metastases. The stage is valuable both to decide treatment for individual patients based on guidelines,

and to define outcomes as a basis for population-level analysis of health programs [17]. For many reasons, in any case, formal staging data is not regularly collected for all cancer patients. Staging of patients is suggested as a standard of care by national cancer bodies, e.g. [18], and provides the base for international benchmarking of outcomes. At the present time because of increasing costs and constrained resources from one viewpoint and significant costs of avoidance, screening and treatment of chronic diseases, particularly cancer on the other hand, Health care providers are searching for the most effective and cost-effective care [19]. While most lung nodules are benign, some can characterize malignancy. Whenever detected early and suitable care rendered, there is the possibility to save many lives and meaningfully decrease costs [20]. Significant clinical information is generally recorded in unstructured free text, and converting it to a structured format can be a time consuming task that may not effectively capture all aspects of the information. Nevertheless, there are at least two large motivations for converting unstructured data into structured data. First, the decrease of time required for manual expert review and second the secondary use of these information for large scale automated processing [20]. Text mining includes a range of approaches used for describing and transforming text [21]. NLP is a collection of syntactic and/or semantic rule- or statistical-based processing algorithms that can be used to parse, segment, extract, or analyze text data in text mining [22]. The use of text analysis differentiates these methods; text mining uses the words as a unit of analysis (e.g., frequencies, the presence or absence of specific words of interest), while NLP techniques use the hidden metadata including content and phrase patterns. Both NLP and text mining have been used in health-related online domains such as mental health [23], oncology [24], and infectious disease [25]. During the previous decade, automatic text processing has been used effectively to extract data from electronic radiology and other reports for a variety of purposes.

Clinical text has some properties such as poor structure, abundant shorthand and domain specific vocabularies that make the application of NLP challenging. Other challenges in the field, like identification of temporal associations, assessment of context-dependent text, and concept normalization to particular terminologies, that remain open [26-28]. We have conducted a systematic literature review to detect the used NLP solutions. However, to the best of our knowledge, no previous systematic literature review has summarized these studies. The aim of this study is to systematically review and assess the reporting and methodological quality of all development which were aimed to diagnosing or staging lung cancer using NLP solutions.

MATERIAL AND METHODS

Eligibility criteria

We followed international Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines in conducting this review [29]. Study characteristics for eligibility were: published research of experimental or methodological studies, conference papers, written in English, using NLP techniques on free text for this review. Editorials, commentaries, letters, books, and presentations were excluded. Publications were excluded if no full text could be retrieved. The application of NLP techniques should be in the field of human healthcare. The source of free text should radiology reports, the NLP technique should relate to diagnose of lung cancer or determining the level of this cancer.

Information sources and searches

Articles were identified for this review through a search of PubMed, Scopus, Web of Science, Embase databases up to Dec 2019. The keywords used to retrieve articles from PubMed data base was as following:

```
("NLP"[Title/Abstract] OR "natural language
process"[Title/Abstract]) AND
("Cancer"[Title/Abstract] OR
"neoplasm"[Title/Abstract])
```

Study selection

After removing duplicate studies, two authors independently reviewed titles and abstracts of all identified studies. Fig 1 illustrates the inclusion process. In the second stage, two authors independently assessed the full-text articles. The disagreements were resolved through discussion and if required with referral to a third researcher. During the screening of the studies, the reviewers documented the reasons for the exclusion of each study. We used a free web and mobile application platform for paper screening (Rayyan QCRI systematic review software) [30]. Included studies were those that: (a) used NLP (b) used radiology reports (c) relate to diagnose of lung cancer or determining level of this cancer.

Data collection process

We developed a customized data-extraction form to be used by 2 reviewers to extract specific details about each study. This form consisted of study location, NLP technique type, sample size, performance, library/platform, and evaluation method (Table 1).

Risk of bias assessment

As various databases were searched to extract the

relevant papers, the evidence selection bias may not occur. To decrease the risk of bias in data extraction a reliable data extraction form was developed and before data extraction the reviewers agreed and discussed the concept of each data items in the form.

Fig 1: Study identification and selection process.

RESULTS

Literature searches identified 231 potentially relevant citations. Of those, 18 studies were deemed eligible for full-text review, and 7 studies fulfilled the inclusion criteria for this systematic review. Table 1 illustrates the characteristics of included studies.

Performance Measures

Performance measures that have been used in the studies in this review include sensitivity (also called recall in the field of NLP), specificity, positive predictive value (PPV) (also called precision), F score (harmonized average of recall and precision) [31] and accuracy. When describing the performance of a method, we focus on sensitivity and specificity; PPV is referred to if specificity is unavailable. Some studies provide the F score, which is frequently used in the field of NLP as a single, overall measure of system performance. All studies reported various performance measures based on the difference between combinations of methods. Most (71%) of studies have reported sensitivity and specificity of the NLP algorithm for identifying the presence of lung nodule(s).

Data Representation and Data Analysis

Three studies developed a NLP algorithm to scan radiology notes and determine the presence or

absence of nodules to identify (all) patients with incident lung nodules for treatment or follow-up. Two studies [17, 32] used natural language processing to transform the report text, including identification of UMLS terms and detection of negated findings to classifying reports, also one of them [17] used an SVM-based text classification system for staging lung cancer patients.

Due to the heterogeneity and multidisciplinary nature of the included studies, a formal meta-analysis was not possible. If a study tested multiple methods for a particular application, we report the performance of the best performing method.

Only two latest studies [32, 33] use deep learning-based approach for extracting information from lung cancer patients' lung CT radiology report

NLP technique used in included studies grouped in 3 categories: rule base (n=2) [32, 33], machine learning (n=3) [17, 32, 33], and combinations of rule base and machine learning (n=2) [34, 35].

Most studies (71%) reported outcomes for only one NLP technique. The most frequent models developed used variants of Machine Learning Models, including SVM and deep learning.

The majority of studies [32, 35-37] used a manually annotated corpus of free-text documents as the comparator. Internal validations in the studies were done using stratified random sampling and 10-fold or n-fold cross-validation.

DISCUSSION

Principal finding

Our systematic review has shown that despite the abundance of studies [32, 34-37] using image processing techniques for lung cancer-related applications, there are only seven studies were done about the automatic extraction of useful information from lung cancer patients' lung CT radiology reports. Although the recent trend of deep learning libraries, toolkits (e.g. TensorFlow, Keras, Lasagne library, deeplearning4j), and studies for natural language processing applications, we found only two latest studies use deep learning-based approach for extracting information from lung cancer patients' lung CT radiology report [33, 38-40]. Probably because deep learning for natural language processing is a developing area of study, most of the reports published in different type of publications like conferences, and electronic preprints repository like arXiv instead of journals.

None of the reviewed studies used public data which shows the scarcity of public data for this particular subject. This is due to the importance of privacy of clinical data, legal and regulatory compliance such as federal law of Health Insurance Portability and

Accountability Act of 1996 (HIPAA), and the Directive 95/46/EC of the European Parliament for data protection. There is a need for shared public data with compliance with regulation to facilitate the

contribution of researchers around the world for improvement of NLP algorithms for clinical and healthcare purposes.

Table 1: Characteristics of included studies

Reference	Authors	Publication Year	Study location	Sample size	NLP type	library/ platform	Evaluation	Performance
[36]	Danforth et al	2012	Kaiser Permanente Southern California	12,389	Rule Based	-	Internal validation: A stratified random sample of 116 patients were used for evaluation.	Sensitivity: 96%(95% CI: 88–99%), specificity: 86%(95% CI: 75–93%), positive predictive value: 87%(95% CI: 77–93%), negative predictive value: 96%(95% CI: 87–99%)
[37]	Beyer et al	2017	USA- a Lung Cancer Screening Center	1501	Rule Based	UIMA	Internal validation: 498 reports were used for evaluation.	Sensitivity (suspicious case) : 74.1% Specificity (suspicious case): 98.6% Sensitivity (positive) : 75 % Specificity (positive): 98.8%
[17]	Iain McCowan et al	2006	Queensland Cancer Control and Analysis Team (QCCAT) in Queensland Health	718	Machine Learning	SVM light	Internal validation: Reports from more than 700 lung cancer patients were analyzed.	T staging (Primary Tumor) sensitivity : 0.61 specificity : 0.88 PPV : 0.62 F1: 0.61 N staging (Regional Lymph Nodes) sensitivity : 0.78 specificity : 0.90 PPV : 0.78 F1: 0.78
[34]	Bipin Karunakaran et al	2017	• Mayo Clinic part-of-speech corpus • GENIA (literature mining project in molecular biology from the University of Tokyo) • Penn Treebank	300000 Validation notes:1096	Rule Based and Machine Learning	cTAKES	Internal validation: A total of 1,096 notes were validated	Precision :0.873 Recall :0.947 Accuracy :0.912 F1: 0.908
[35]	Roxanne Wadia et al	2018		326	Rule Based and Machine Learning	Yale cTAKES	Internal validation: A total of 450 reports representing 428 patients were analyzed	NLP sensitivity: 77.3% manual coding sensitivity: 51.5% NLP specificity: 72.5% manual coding specificity: 82.5% NLP PPV: 88.4% manual coding PPV :88.9% NLP NPV: 54% manual coding NPV: 38.5%
[32]	Liwei Wang et al	2019	Mayo Clinic EHR at Rochester site	2311	Machine Learning	-	Internal validation: precision range from .704 to 1, Recall from .779 to 1 for 30 days to 365 days	recall stage:89 recall histology: 98 recall tumor grade: 78 recall therapies achieving: 100 precision stage:70 precision histology: 88 precision tumor grade: 90 precision therapies achieving: 100
[33]	Er. Khushbu Gupta et al	2019	The Department of Radiology, Ramathibodi Hospital, Mahidol University	2000 CT-scan reports of lung cancer patients	Machine Learning	-	Internal validation: accuracy 59% for T-staging, 36% for N-staging and 41% for M-staging	accuracy up to 85%

Due to difficulty in the acquisition of radiology reports from other institutions most authors would be forced to use data from the same institution for train and validation. This will result in a lack of external validations or scarcity of external validation

in NLP studies in medicine, as such external validation was absent from all of the included studies of this review. Implementation of NLP methods in a real environment is also necessary. Despite commercially available NLP tools for radiology

reports, none of the included studies has implemented in healthcare institutions. Lack of generalizability due to paucity of data for external validation and clinicians' distrust due to the unclear nature of the way of generating output evidently are important contributing factors to being not popular in practice.

In this review, we have only included application of NLP methods for lung cancer radiology report, but NLP techniques have been adapted in other kinds of radiology reports [41-43]. There are many other NLP medical applications such as EHR information extraction, and identification of diseases [44-46]. The Use of NLP in radiology report applications may benefit from NLP applications that operate on EHR data from other fields. For instance, the NLP task of identifying a clinical diagnosis from broader EHR content can help provide a more definitive reference standard to diagnoses.

Limitations

To get quality studies, we searched four prestigious academic research databases (Pubmed, Scopus, Web of Science, and Embase). Because of it, we may miss some new studies. Conference and electronic preprint manuscripts are more likely to be biased, which will lead to misrepresenting performance parameters to excluding studies that are more probable to be biased justify our decision.

In this review we found studies for identification of cancer patients, finding most urgent suspected lung cancer cases for follow up, determine the probability of having lung cancer, and determine the stage of lung cancer, in which various NLP methods were applied. The diversity of the different situations prevents us to do a meta-analysis of the studies. Also, there was only one study that compared two different NLP methods, and hence it is difficult to conclude on which NLP techniques is the best.

Recommendations

Consider the great potential of deep learning approaches for natural language processing, and the benefits of high accuracy automated information extraction of lung cancer patients' chest CT radiology reports, there is a need for more research in this area.

Most of NLP applications in cancer diagnosis and treatment remain in proof of concept stage and never merge in clinical practice routine. Identify the existing challenges and finding their solutions help clinicians to save more lives at the end of days.

CONCLUSION

Evaluation of studies in diagnosing and staging methods in lung cancer using natural language processing (NLP) shows there is a number of studies on diagnosing lung cancer but there are a few works on staging that. In some studies, combination of methods was considered and NLP isolated was not sufficient for capturing satisfying results. There are potentials to improve studies by adding other data sources, further refinement and subsequent validation.

AUTHOR'S CONTRIBUTION

The authors agree on this final form of the manuscript, and attested that all authors contributed in the final draft of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

FINANCIAL DISCLOSURE

No financial interests related to the material of this manuscript have been declared.

REFERENCES

1. Baudendistel I, Winkler E, Kamradt M, Brophy S, Längst G, Eckrich F, et al. Cross-sectoral cancer care: Views from patients and health care professionals regarding a personal electronic health record. *Eur J Cancer Care (Engl)*. 2017; 26(2): e12429. PMID: 26840784 DOI: 10.1111/ecc.12429 [[PubMed](#)]
2. Azami-Aghdash S, Ghojzadeh M, Sheyklo SG, Daemi A, Kolahdouzan K, Mohseni M, et al. Breast cancer screening barriers from the womans perspective: A meta-synthesis. *Asian Pac J Cancer Prev*. 2015; 16(8): 3463-71. PMID: 25921163 DOI: 10.7314/apjcp.2015.16.8.3463 [[PubMed](#)]
3. Murphy KM, Nguyen V, Shin K, Sebastian-Deutsch A, Frieden L. Health care professionals and the employment-related needs of cancer survivors. *J Occup Rehabil*. 2017; 27(2): 296-305. PMID: 26857027 DOI: 10.1007/s10926-016-9629-2 [[PubMed](#)]
4. Maleki D, Ghojzadeh M, Mahmoudi S-S, Mahmoudi S-M, Pournaghi-Azar F, Torab A, et al. Epidemiology of oral cancer in Iran: A systematic review. *Asian Pac J Cancer Prev*. 2015; 16(13): 5427-32. PMID: 26225689 DOI: 10.7314/apjcp.2015.16.13.5427 [[PubMed](#)]
5. Ferraz AFB, Rosim R, Anaya P. Standardization process of raw datasus and consumption analysis of oncology therapies in the Brazil public health care system: A comparison between raw and standardized dataset in colorectal and lung cancer. *Value in Health*. 2015; 18(7): A811.
6. Faris N, Yu X, Sareen S, Signore RS, McHugh LM, Roark K, et al. Preoperative evaluation of lung cancer in a community health care setting. *Ann Thorac Surg*. 2015; 100(2): 394-400. PMID: 26074001 DOI: 10.1016/j.athoracsurg.2015.03.008 [[PubMed](#)]

7. Ryoo JJ, Malin JL, Ordin DL, Oishi SM, Kim B, Asch SM, et al. Facility characteristics and quality of lung cancer care in an integrated health care system. *J Thorac Oncol.* 2014; 9(4): 447-55. PMID: 24736065 DOI: 10.1097/JTO.000000000000108 [[PubMed](#)]
8. Sundaram B, Kazerooni E. Preface: Lung cancer is an important public health care issue. *Radiol Clin North Am.* 2012; 50(5): xi. PMID: 22974784 DOI: 10.1016/j.rcl.2012.06.009 [[PubMed](#)]
9. Torre LA, Siegel RL, Jemal A. Lung cancer statistics. In: Ahmad A, Gadgeel S (eds). *Lung cancer and personalized medicine*. Springer; 2016.
10. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019; 69(1): 7-34. PMID: 30620402 DOI: 10.3322/caac.21551 [[PubMed](#)]
11. Hunnibell LS, Rose MG, Connery DM, Grens CE, Hampel JM, Rosa M, et al. Using nurse navigation to improve timeliness of lung cancer care at a veterans hospital. *Clin J Oncol Nurs.* 2012; 16(1): 29-36. PMID: 22297004 DOI: 10.1188/12.CJON.29-36 [[PubMed](#)]
12. Hirano H, Maeda H, Yamaguchi T, Yokota S, Mori M, Sakoda S. Survivin expression in lung cancer: Association with smoking, histological types and pathological stages. *Oncol Lett.* 2015; 10(3): 1456-62. PMID: 26622690 DOI: 10.3892/ol.2015.3374 [[PubMed](#)]
13. Huang R, Wei Y, Hung RJ, Liu G, Su L, Zhang R, et al. Associated links among smoking, chronic obstructive pulmonary disease, and small cell lung cancer: A pooled analysis in the International Lung Cancer Consortium. *EBioMedicine.* 2015; 2(11): 1677-85. PMID: 26870794 DOI: 10.1016/j.ebiom.2015.09.031 [[PubMed](#)]
14. Yun YD, Back JH, Ghang H, Jee SH, Kim Y, Lee SM, et al. Hazard ratio of smoking on lung cancer in Korea according to histological type and gender. *Lung.* 2016; 194(2): 281-9. PMID: 26718701 DOI: 10.1007/s00408-015-9836-1 [[PubMed](#)]
15. Thun MJ, Hannan LM, Adams-Campbell LL, Boffetta P, Buring JE, Feskanich D, et al. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS Medicine.* 2008; 5(9): e185. PMID: 18788891 DOI: 10.1371/journal.pmed.0050185 [[PubMed](#)]
16. Barletta JA, Yeap BY, Chirieac LR. Prognostic significance of grading in lung adenocarcinoma. *Cancer.* 2010; 116(3): 659-69. PMID: 20014400 DOI: 10.1002/cncr.24831 [[PubMed](#)]
17. McCowan I, Moore D, Fry M-J. Classification of cancer stage from free-text histology reports. *Conf Proc IEEE Eng Med Biol Soc.* 2006; 2006: 5153-6. PMID: 17945879 DOI: 10.1109/IEMBS.2006.259563 [[PubMed](#)]
18. Lewis SZ, Diekemper R, Addrizzo-Harris DJ. Methodology for development of guidelines for lung cancer: diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest.* 2013; 143(5): 41S-50S. PMID: 23649432 DOI: 10.1378/chest.12-2344 [[PubMed](#)]
19. Azar FE, Azami-Aghdash S, Pournaghi-Azar F, Mazdaki A, Rezapour A, Ebrahimi P, et al. Cost-effectiveness of lung cancer screening and treatment methods: A systematic review of systematic reviews. *BMC Health Serv Res.* 2017; 17(1): 413. PMID: 28629461 DOI: 10.1186/s12913-017-2374-1 [[PubMed](#)]
20. Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform.* 2017; 73: 14-29. PMID: 28729030 DOI: 10.1016/j.jbi.2017.07.012 [[PubMed](#)]
21. Dreisbach C, Koleck TA, Bourne PE, Bakken S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int J Med Inform.* 2019; 125: 37-46. PMID: 30914179 DOI: 10.1016/j.ijmedinf.2019.02.008 [[PubMed](#)]
22. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: An introduction. *J Am Med Inform Assoc.* 2011; 18(5): 544-51. PMID: 21846786 DOI: 10.1136/amiajnl-2011-000464 [[PubMed](#)]
23. Calvo RA, Milne DN, Hussain MS, Christensen H. Natural language processing in mental health applications using non-clinical texts. *Natural Language Engineering.* 2017; 23(5): 649-85.
24. Yim W-W, Yetisgen M, Harris WP, Kwan SW. Natural language processing in oncology: A review. *JAMA Oncol.* 2016; 2(6): 797-804. PMID: 27124593 DOI: 10.1001/jamaoncol.2016.0213 [[PubMed](#)]
25. Paul MJ, Sarker A, Brownstein JS, Nikfarjam A, Scotch M, Smith KL, et al. Social media mining for public health monitoring and surveillance. *Pacific Symposium on Biocomputing*. World Scientific; 2016.
26. Sun W, Rumshisky A, Uzuner O. Evaluating temporal relations in clinical text: 2012 i2b2 Challenge. *J Am Med Inform Assoc.* 2013; 20(5): 806-13. PMID: 23564629 DOI: 10.1136/amiajnl-2013-001628 [[PubMed](#)]
27. Uzuner O, Bodnari A, Shen S, Forbush T, Pestian J, South BR. Evaluating the state of the art in coreference resolution for electronic medical records. *J Am Med Inform Assoc.* 2012; 19(5): 786-91. PMID: 22366294 DOI: 10.1136/amiajnl-2011-000784 [[PubMed](#)]
28. Elhadad N, Pradhan S, Gorman S, Manandhar S, Chapman W, Savova G. SemEval-2015 task 14: Analysis of clinical text. *International Workshop on Semantic Evaluation. Association for Computational Linguistics*; 2015.
29. Moher D, Shamseer L, Clarke M, Ghera D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015; 4(1): 1. PMID: 25554246 DOI: 10.1186/2046-4053-4-1 [[PubMed](#)]
30. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid AJSr. Rayyan: A web and mobile app for systematic reviews. *Syst Rev.* 2016; 5(1): 210. PMID: 27919275 DOI: 10.1186/s13643-016-0384-4 [[PubMed](#)]

31. Van Rijsbergen CJ. Information retrieval. 2nd ed. London, England: Butterworths; 1979.
32. Wang L, Luo L, Wang Y, Wampfler J, Yang P, Liu H. Natural language processing for populating lung cancer clinical research data. *BMC Med Inform Decis Mak*. 2019; 19(Suppl 5): 239. PMID: 31801515 DOI: 10.1186/s12911-019-0931-8 [[PubMed](#)]
33. Gupta EK, Thammasudjarit R, Thakkinstian A. NLP automation to read radiological reports to detect the stage of cancer among lung cancer patients. Workshop on Widening NLP; 2019.
34. Karunakaran B, Misra D, Marshall K, Mathrawala D, Kethireddy S. Closing the loop: Finding lung cancer patients using NLP. International Conference on Big Data. IEEE; 2017.
35. Wadia R, Akgun K, Brandt C, Fenton BT, Levin W, Marple AH, et al. Comparison of natural language processing and manual coding for the identification of cross-sectional imaging reports suspicious for lung cancer. *JCO Clin Cancer Inform*. 2018; 2: 1-7. PMID: 30652545 DOI: 10.1200/CCI.17.00069 [[PubMed](#)]
36. Danforth KN, Early MI, Ngan S, Kosco AE, Zheng C, Gould MK. Automated identification of patients with pulmonary nodules in an integrated health system using administrative health plan data, radiology reports, and natural language processing. *J Thorac Oncol*. 2012; 7(8): 1257-62. PMID: 22627647 DOI: 10.1097/JTO.0b013e31825bd9f5 [[PubMed](#)]
37. Beyer SE, McKee BJ, Regis SM, McKee AB, Flacke S, El Saadawi G, et al. Automatic Lung-RADS™ classification with a natural language processing system. *J Thorac Dis*. 2017; 9(9): 3114-22. PMID: 29221286 DOI: 10.21037/jtd.2017.08.13 [[PubMed](#)]
38. Attardi G. DeepNL: A deep learning NLP pipeline. Workshop on Vector Space Modeling for Natural Language Processing. Association for Computational Linguistics; 2015.
39. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, McClosky D. The Stanford core NLP toolkit. Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics; 2014.
40. Chen MC, Ball RL, Yang L, Moradzadeh N, Chapman BE, Larson DB, et al. Deep learning to classify radiology free-text reports. *Radiology*. 2017; 286(3): 845-52. PMID: 29135365 DOI: 10.1148/radiol.2017171115 [[PubMed](#)]
41. Alex B, Grover C, Tobin R, Sudlow C, Mair G, Whiteley W. Text mining brain imaging reports. *J Biomed Semantics*. 2019; 10(Suppl 1): 23. PMID: 31711539 DOI: 10.1186/s13326-019-0211-7 [[PubMed](#)]
42. Viani N, Miller TA, Napolitano C, Priori SG, Savova GK, Bellazzi R, et al. Supervised methods to extract clinical events from cardiology reports in Italian. *J Biomed Inform*. 2019; 95: 103219. PMID: 31150777 DOI: 10.1016/j.jbi.2019.103219 [[PubMed](#)]
43. Liu Y, Liu Q, Han C, Zhang X, Wang X. The implementation of natural language processing to extract index lesions from breast magnetic resonance imaging reports. *BMC Med Inform Decis Mak*. 2019; 19(1): 288. PMID: 31888615 DOI: 10.1186/s12911-019-0997-3 [[PubMed](#)]
44. Viani N, Kam J, Yin L, Bittar A, Dutta R, Patel R, et al. Temporal information extraction from mental health records to identify duration of untreated psychosis. *J Biomed Semantics*. 2020; 11(1): 2. PMID: 32156302 DOI: 10.1186/s13326-020-00220-2 [[PubMed](#)]
45. Chandran D, Robbins DA, Chang C-K, Shetty H, Sanyal J, Downs J, et al. Use of natural language processing to identify obsessive compulsive symptoms in patients with schizophrenia, schizoaffective disorder or bipolar disorder. *Sci Rep*. 2019; 9(1): 14146. PMID: 31578348 DOI: 10.1038/s41598-019-49165-2 [[PubMed](#)]
46. Savova GK, Danciu I, Alamudun F, Miller T, Lin C, Bitterman DS, et al. Use of natural language processing to extract clinical cancer phenotypes from electronic medical records. *Cancer Res*. 2019; 79(21): 5463-70. PMID: 31395609 DOI: 10.1158/0008-5472.CAN-19-0579 [[PubMed](#)]