

The Growing Impact of Natural Language Processing in Healthcare and Public Health

INQUIRY: The Journal of Health Care
Organization, Provision, and Financing
Volume 61: 1–10
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00469580241290095
journals.sagepub.com/home/inq



Aadit Jerfy¹, Owen Selden¹, and Rajesh Balkrishnan, PhD¹ 

Abstract

Natural Language Processing (NLP) is a subset of Artificial Intelligence, specifically focused on understanding and generating human language. NLP technologies are becoming more prevalent in healthcare and hold potential solutions to current problems. Some examples of existing and future uses include: public sentiment analysis in relation to health policies, electronic health record (EHR) screening, use of speech to text technology for extracting EHR data from point of care, patient communications, accelerated identification of eligible clinical trial candidates through automated searches and access of health data to assist in informed treatment decisions. This narrative review aims to summarize the current uses of NLP in healthcare, highlight successful implementation of computational linguistics-based approaches, and identify gaps, limitations, and emerging trends within the subfield of NLP in public health. The online databases Google Scholar and PubMed were scanned for papers published between 2018 and 2023. Keywords “Natural Language Processing, Health Policy, Large Language Models” were utilized in the initial search. Then, papers were limited to those written in English. Each of the 27 selected papers was subject to careful analysis, and their relevance in relation to NLP and healthcare respectively is utilized in this review. NLP and deep learning technologies scan large datasets, extracting valuable insights in various realms. This is especially significant in healthcare where huge amounts of data exist in the form of unstructured text. Automating labor intensive and tedious tasks with language processing algorithms, using text analytics systems and machine learning to analyze social media data and extracting insights from unstructured data allows for better public sentiment analysis, enhancement of risk prediction models, improved patient communication, and informed treatment decisions. In the recent past, some studies have applied NLP tools to social media posts to evaluate public sentiment regarding COVID-19 vaccine use. Social media data also has the capacity to be harnessed to develop pandemic prediction models based on reported symptoms. Furthermore, NLP has the potential to enhance healthcare delivery across the globe. Advanced language processing techniques such as Speech Recognition (SR) and Natural Language Understanding (NLU) tools can help overcome linguistic barriers and facilitate efficient communication between patients and healthcare providers.

Keywords

health informatics, artificial intelligence, natural language, healthcare management, managed care

What do we already know about this topic?

Natural Language Processing (NLP) technologies are becoming more prevalent in healthcare and hold potential solutions to current challenges (eg.: electronic health record screening, gaps in medical record data, patient communications).

How does your research contribute to the field?

Our narrative review summarizes the current uses of NLP in healthcare, highlight successful implementation of computational linguistics-based approaches, and identify gaps, limitations, and emerging trends within the subfield of NLP in public health and healthcare.

What are your research's implications toward theory, practice, or policy?

Continued development of NLP applications in healthcare will contribute to several improvements in areas such as public sentiment analysis, enhancement of risk prediction models, and improved patient communication in the future.



Introduction

Computational Linguistics/Natural Language Processing Overview

Computational Linguistics (CL) and Natural Language Processing (NLP) are fields at the intersection of computer science and linguistics. CL is defined as the development and application of computational models to understand, analyze, and process human language. NLP is a subfield of computational linguistics, focused on Artificial Intelligence (AI) models that interpret and generate human language. NLP utilizes AI-based techniques to extract information from large text samples, derive its meaning, and respond to human inputs. Both play an important role in advancing methods by which computers interact with human language.

Objectives

The primary objective of this paper is to examine the role of Computational Linguistics (CL) in addressing critical challenges in the fields of public health and healthcare. By conducting an in-depth analysis of existing literature, this review will summarize the current uses of NLP applications in healthcare. This review will also highlight successful implementation of computational linguistics-based approaches. Gaps, limitations, and emerging trends within this sub-field will be addressed, allowing for a comprehensive analysis of NLP's potential integration into healthcare. This review will provide analysis of the solutions that CL and NLP provide in public health advertising, medical record analysis, and patient communication. The broader goal of our review is to increase accessibility of information that will inform future research, policy, and implementation of NLP/CL techniques in healthcare.

Relevance

NLP and CL offer novel opportunities for medical professionals to analyze massive amounts of unstructured data (eg, clinical notes¹, medical records, public health literature, public reaction to health policies on social media), facilitating valuable insights. "80% of medical data remains unstructured and untapped after it is created (eg, text, image, signal, etc.) (Figure 1) Since it is hard to handle this type of data, it tends to be ignored, unsaved, or abandoned in most medical centers for a long time."¹⁸ As the digital era continues to advance the methods by which information can be generated,

linguistic and computational tools have the potential to enhance current practices. NLP technology has the potential to enhance the efficacy of diagnostics and suggest potential treatment and patient care plans. These examples showcase only a portion of the potential impact that NLP could have in public health and related fields. By analyzing the connections between computation, language, and healthcare, this paper provides a synopsis of the intersection of computational linguistics and medicine.

Methods

Information Sources

The material utilized in this narrative review was sourced from the online databases Google Scholar and PubMed (as they have the largest scope and are most relevant to NLP), through searches for papers published between 2018 and 2023. The following keywords were utilized: natural language processing, health policy, large language models. The search criteria was limited to papers written in English with a focus on healthcare and advertising. A brief disclaimer: the company now known as "X" is referred to as Twitter throughout the paper, as the studies examined were published before the rebranding of the company.

Abstraction and Evaluation

This literature review found 49 articles meeting the study criteria keywords during the study time frame gathered through searches on PubMed and Google Scholar. All irrelevant results (unrelated to healthcare, NLP, and public opinion) and duplicate articles were removed, leaving $n = 27$ papers¹⁻²⁷. These articles all met the eligibility requirements for this review, allowing for their inclusion.

Trends/Potential Uses

Analysis of Public Opinion on Health Policies

NLP technologies, especially following the COVID-19 pandemic, have been utilized to analyze public opinion on various health policies. COVID-19, in particular, demonstrated significant variation in public opinion with regional variation in vaccination and mask policies. An international, comprehensive study conducted by the University of Utah over Reddit servers in Australia, Canada, the United States, and the United Kingdom, delved into opinions posted on forums and

¹University of Virginia School of Medicine, Charlottesville, VA, USA

Received 1 June 2024; revised 10 September 2024; revised manuscript accepted 18 September 2024

Corresponding Author:

Rajesh Balkrishnan, Department of Public Health Sciences, University of Virginia School of Medicine, P.O. Box 800717 Charlottesville, VA 22908 USA.
Email: rb9ap@virginia.edu

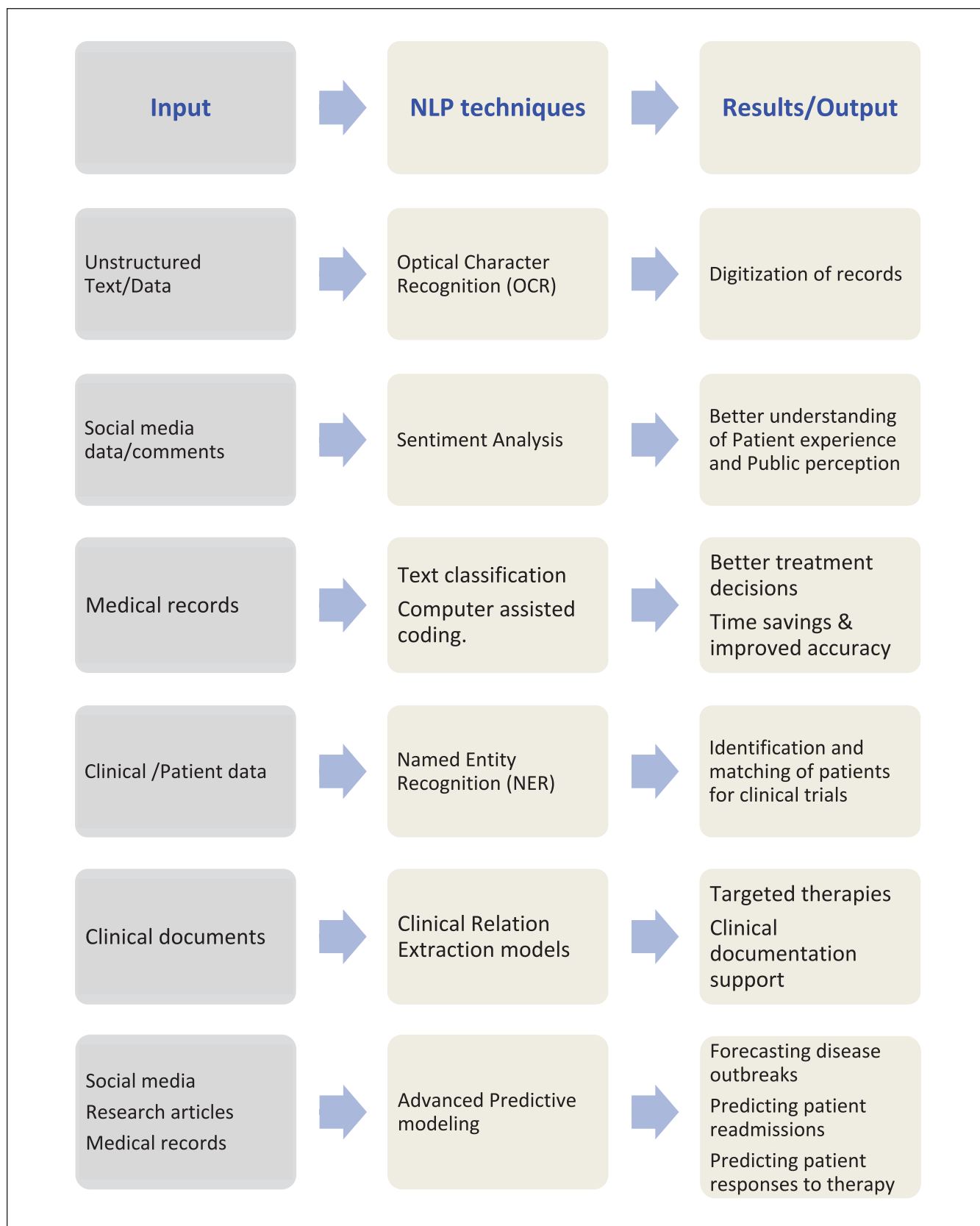


Figure 1. Some of the potential advances enabled by application of NLP techniques and lists some of the tools used.

discussions, allowing for insights into what these users were feeling.¹³ Over 84 000 posts and 1 000 000 comments (Feb 2020–Nov 2020) were explored by latent dirichlet allocation (LDA) software. LDA is a common technique in NLP that allows for the inference of topics from large bodies of text. Important distinctions in the discussion of similar topics between the regions on Reddit were revealed. The study showed that an NLP based approach could offer deeper perspectives than survey-based approaches. NLP was able to do so because of the sheer amount of data able to be screened. A similar study was funded by the Japan Science and Technology Agency, where NLP was utilized to learn more about public views on the COVID vaccine.²³ Using an NLP model, 144 101 tweets from Japan (August 2020–June 2021) were reviewed, helping professionals identify relationships between public views and infection, death, and vaccination levels. The LDA model identified roughly 85% of the tweets to express neutral opinions on vaccination, with negative opinions outweighing positive for the remaining 15%. Keywords such as “receive, side effect, prevention, danger, government, infect,” allowed the model to separate the neutral, negative, and positive reviews cited above. The NLP model, in this study, allowed for screening and sorting of massive amounts of data, which would have been less efficient if done by human experts. There are numerous studies supporting this claim, particularly in relation to COVID-19 and Table 1 depicts the data analyzed by similar methods, and results/conclusions of each.

Collectively, the above studies showcase NLP’s existing use in analysis of public opinion on health policy, being able to analyze millions of data points across these cases during a pandemic when so many people utilized social media to convey their thoughts. Across the majority of these studies, trends relating to both positive and negative posts were correlated with rates of infection and death, along with efficacy of various vaccines and their development. Identification of thematic trends in social media posts tracks public opinion and allows for in-depth responses from public health officials that achieve higher satisfaction and compliance.

These uses of LDA, topic modeling, and methods such as unsupervised ML, coupled with sentiment analysis, depict NLP’s adaptability in various research contexts. Through analysis of papers using these methods, insights into public response to realms beyond COVID-19 policy can be made, although COVID-19 response contains the most prominent examples, as the pandemic provided a need for NLP technology capable of such large-scale analysis.

Electronic Health Record Analysis/Risk Prediction

Natural language processing can be utilized for analysis of various data types, including textual clinical data and patient electronic health records. Manual review of patient records is a tedious process. Trends and similarities in clinical texts correlate with risk of future medical complications and hospitalization. With NLP of these textual bodies, predictive

models can be created, scanning patient clinical data and forecasting admission into medical facilities.

Across various studies, NLP methods have consistently proven their value in enhancing predictive performance. One such study delves into the incorporation of NLP to construct a comprehensive analytical framework.²² Focusing on accurate prediction of mortality outcomes in ICU patients, the combination of NLP derived keywords and terms consistently enhanced model performance and increased the area under the receiver operating characteristic curve (AUC) from 0.831 to 0.922 (AUC is a measure of predictive accuracy with values ranging from 0 to 1. A higher value indicates higher predictive accuracy).²⁷ Such a significant jump in AUC shows the value of NLP in predictive risk modeling, a tendency not exclusive to ICU mortality models.

In an NLP model designed to identify risk factors correlating with lower back pain from the electronic medical record of patients, only 347 out of the 2749 health records of patients exhibiting back pain sampled contained one or more of the keywords related to said risks.¹⁵ Despite downsizing, the input dataset still resulted in a 2% increase in the AUC. While this sort of sampling may result in biases of the model in analysis—being more likely to identify text similar to its limited training data—the fact that AUC was able to improve in these circumstances shows the promise of the NLP techniques utilized in this case.

Another example of natural language processing of medical records can be found in a study relating to gender identity.¹¹ Precise gender identity identification is essential in ensuring equitable healthcare, particularly within gender-diverse populations. Electronic Health Records (EHR) fields often contain incomplete or insufficient gender information; through use of gender-related keywords and implementation of NLP and deep learning, the group constructed an accurate predictive model for patient gender identity. This proved that NLP models, integrated with deep learning techniques, can be utilized for efficient prediction of patient gender identity when trained on limited EHR data.

Initial scanning techniques in this subfield are often generalizable, as the processes for large-scale textual file analysis tend to be fairly consistent across various applications of it. Similar methods were utilized in a study of 76 479 patients over age 65 with multimorbidity (more than one incurable illness), where NLP was used to extract psychosocial factors (eg, isolation, housing, financial insecurity, stress) from chart notes, and predict their impact on 1-year hospitalization. Other variables were isolated from the EHR of said patients, and logistic regression was used to predict the likelihood of health issues in relation to these factors. The model found that patients with any of the aforementioned psychosocial factors had higher baseline utilization of healthcare facilities, and a greater number of chronic illnesses. Predictive models that did not evaluate with consideration to psychosocial factors performed worse, and had overall less efficient discrimination. NLP of tens of thousands of chart notes allowed for

Table 1. Summary of studies related to COVID-19 identified in the literature review.

Lead author	Study	Relevance	Main Findings
Hu M	Perspectives of the COVID-19 Pandemic on Reddit: Comparative Natural Language Processing Study of the United States, the United Kingdom, Canada, and Australia	The study was aimed at comparing public attitudes toward COVID-19. The Reddit platform was used as a data source. Topic modeling NLP methods (LDA) were used for investigation and comparison.	Across all four countries included in the study, there was a consistent decline in the volume of postings over the timeframe examined. Lockdown events corresponded to peaks in volumes. Subreddits from the United Kingdom and Australia contained more evidence-based policy discussion as compared to those from the United States and Canada.
Al-Ahdal T	Improving Public Health Policy by Comparing the Public Response during the Start of COVID-19 and Monkeypox on Twitter in Germany: A Mixed Methods Study	The intent was to capture public sentiment and document problems and prescribe solutions. Machine learning was combined with sequential explanatory mixed methods.	A mixed methods design allowed the addition of context to machine learning results. Results showed very different public responses to each of the crises. During monkeypox, the public expressed minimal regard for precautionary actions. In the case of COVID-19, the focus was more local. As death rates increased, protective measures impacted more and more people. Monkeypox cases had a more global focus. Local community issues were considered minor. Deaths were not a major concern in this case. Findings suggest that tools intended to monitor popular sentiment and response to health issues could be leveraged to promote various policies.
Niu Q	Public Opinion and Sentiment Before and at the Beginning of COVID-19 Vaccinations in Japan: Twitter Analysis	Initial vaccination in Japan faced slow progress. This study aims to comprehend the forces influencing vaccine hesitancy, and generate strategies for further promotion of vaccines.	Both positive and negative sentiments could be triggered by the same event. The LDA model of early-stage negative tweets suggested that factors related to the slow vaccination process may have been at play, thereby highlighting the need for swift policy action. Strong concerns regarding vaccine safety underlined the need for building vaccine confidence.
Hussain Z	Artificial Intelligence-Enabled Social Media Analysis for Pharmacovigilance of COVID-19 Vaccinations in the United Kingdom: Observational Study	This study was aimed at assessing how often and in what manner Adverse Events Following Immunizations (AEFI) appeared in online platforms and offer views into popular sentiments toward vaccines.	There was consistency between the frequently discussed COVID-19 AEFIs found in both governmental sources and those discussed on social media. The analysis also pointed to safety signals detected elsewhere being investigated. The findings supported use of social media analysis for pharmacovigilance purposes; as a complement to conventional knowledge sources.
Liew TM	Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672 133 Twitter Posts	Social media data was utilized to gather current public opinions about COVID-19 vaccines to eventually facilitate vaccination with minimal public dissatisfaction.	The findings may assist in developing plans to increase vaccination rates. Areas requiring greater focus were identified, and based on this information, policies were able to be adjusted. Social media played a part in rates of COVID vaccination.
Huang Z	Public Perception of the Use of Digital Contact-Tracing Tools After the COVID-19 Lockdown: Sentiment Analysis and Opinion Mining	The objective of the study was to grasp levels of public awareness and attitudes regarding utilization of a digital contact tracing tool.	Variations in public opinion on obligatory digital contact tracing were influenced by federal activity and advertising. Customized communication approaches and open handling of data eased fears, and bolstered confidence in the government.

the identification of said factors, and therefore improved model performance.⁸

Audio Recordings

NLP technology has significant potential to address issues within clinical settings as well. Some of these issues were especially evident during the pandemic, particularly involving the identification and understanding of COVID-19 symptoms within clinics.⁵ During the peak of the pandemic healthcare centers were overworked, and often patient symptoms were not accurately transcribed into their records. Any NLP-based analysis of these resulting skewed EHR records would be restricted by the information entered. In addition to this, lab tests were scarcely available, so diagnosis often depended on symptoms observed during clinic visits. These symptoms are diverse, including the common fever, cough, sore throat, and dyspnea. However, there are less common symptoms associated with COVID, including insomnia, rhinorrhea, nausea, myalgias, and many more.

The study proposed to instead examine audio recordings of patient visits, providing a comprehensive and complete source of symptomatic data for NLP technologies to examine. With audio recordings of patient interaction, lesser known symptoms and keywords that might be missed during transcription to the EHR can be recorded. The combination of audio data and EHR data would improve accuracy of categorization for diagnosis, accounting for the limitations of the health record. In addition to audio recordings of patient interactions, the authors of this study seek to acquire data from telehealth appointments, as well as from voice-activated systems including the Amazon Alexa, Google Home, and Siri. With audio transcriptions being a primary source of data, there are ample repositories to pull from to train predictive models with presumably high accuracy (due to the sheer amount of available data).

While there are limitations to such methods, including the need for diligent protection of patient information and complications involved with accessing existing patient recordings, audio transcription-based NLP allows for data analysis on a greater scale than traditional methods.

Patient Communication

NLP presents significant potential in the evolution of patient communication going forward. It enables the analysis and understanding of complex language for improved interactions between healthcare providers and patients. Large Language Models (LLMs) are NLP models with advanced comprehension of human language, being trained on massive amounts of textual data and containing over 100 billion parameters. These models can be trained to understand and simplify the complex language associated with doctor-patient communications.

A study published in the *Yale Journal of Biology and Medicine* describes how electronic health information (eg, imaging reports) has become more accessible to patients due to the 21st Century Cures Act, passed in December of 2016—which, aside from speeding up various regulatory processes, “improved patient access to their electronic health information,” including imaging reports.³ Despite the benefits of this act, these reports are complex, being intended for internal communication between doctors. They are often specific and detailed, with medical terms typically beyond the comprehension of most patients. Patient understanding of their health is crucial in improving future treatment, decreasing hospitalizations, and improving long term wellness; this lack of patient comprehension and gap in communication undoubtedly has a negative affect on future health outcomes.

There are proposed solutions to this problem, including summary statements, second reports, and providing the radiologist’s phone number. However, these all complicate the radiologist’s tasks, distracting them from already strenuous daily requirements. LLMs such as ChatGPT by OpenAI—alongside similar deep learning models including Google’s Bard AI and Microsoft Bing—driven by NLP have the ability to simplify imaging reports without causing the aforementioned disruptions for the radiologist. Approaches incorporating DL-NLP have already been utilized in “identifying critical findings, categorizing oncologic responses, finding follow-up recommendations, identifying pulmonary emboli, detecting complications of stroke, and classifying epilepsy brain MRIs.”³ Between the numerous studies conducted in this area, LLMs show potential in report simplification for patient communications, though expert verification of the output is still necessary before large-scale implementation.

Discussion and Conclusions

Public Sentiment

The various studies that showcase NLP technologies’ roles in mass analysis of public opinion on health policies are linked in both conclusions and methodology. They depict NLP’s ability to process large volumes of data from public forums such as Reddit and Twitter, and uncover differing sentiments from across the globe. LDA techniques, topic modeling, unsupervised machine learning, were coupled with sentiment analysis in these studies, allowing for efficient screening and sorting of the vast datasets provided by these social media sources. Despite these evident advantages, there are risks associated with sole reliance on NLP analyses. Unrepresentative keyword input, or bias in training datasets can distort the outcomes of predictive models. These studies emphasize the need for algorithm refinement to eliminate biases. A balance between automated predictive models and expert verification would ensure reliability of predictive outputs. Transparency in methodology,

frequent algorithmic performance evaluation, and collaboration between researchers is necessary to increase the accuracy and consistency of NLP-based models, especially when data found on social media platforms is inherently subjective.

EHR Analysis

NLP could facilitate the manual review of patient and electronic health records is an unnecessarily time-consuming process. Just as we extract insights from textual bodies and symptoms, NLP is able to scan for keywords associated with various diseases. Predictive risk assessment models can be trained based on this NLP keyword detection, with greater accuracy than standard neural networks. While the advantages of NLP based deep learning approaches are numerous, training based on EHR datasets poses issues in relation to patient data and confidentiality. Vulnerabilities in NLP models could result in unintentional data breaches, potentially compromising patient information. One potential solution to address this moving forward is the incorporation of blockchain technology, which would provide security for any EHR utilized. In simple terms, blockchain is a form of storage, weaving secured data into interconnected sequences.²⁸ Blockchain networks contain a variety of cryptographic techniques and consensus mechanisms in order to ensure data integrity. Patient information can be encrypted and stored, and with the use of randomly generated private keys being required for access, confidentiality will be preserved. This technology would mitigate many vulnerabilities associated with models utilizing electronic health records, resulting in enhanced security for trials involving them.

Audio Transcription to EHR

Despite the advantages associated with EHR analysis, these records are frequently imperfect, with potential keywords associated with specific diseases not being entered. However, audio recordings, and subsequent transcriptions of patient interactions with doctors, eliminate this issue. If EHR entries are made with a complete collection of the keywords mentioned by patients and doctors during appointments, predictive models utilizing NLP will be able to function at a higher level. Records containing information relevant to specific diseases would be able to be identified via NLP, allowing for the patient's risks to be properly assessed. However, similar challenges arise with audio recordings as do with any EHR analysis. Patient confidentiality must be protected, and advanced security measures must be taken to do so. In addition to this, patient consent in being recorded in the first place is necessary, and measures would have to be taken to be as transparent as possible with every patient whose health records would be affected by such methods.

Patient Communication

As highlighted in the review, electronic health reports are becoming more and more accessible to patients. While access to these reports is important, understanding them can be quite complicated, as they are abstruse in nature. Imaging reports are perfect examples of this; originally meant for internal communication within hospitals and between doctors, terms utilized in them would be lost on the majority of patients. LLMs such as ChatGPT can simplify this complex language into phrases and expressions understandable by the average person. Overall ease of patient communication would be heightened through the utilization of LLMs, as general understanding of specific health conditions and treatments would also be increased. Though, as with any AI model, expert verification is necessary before simplified "translations" are provided for patients. Accuracy must be a priority when using LLMs in such a crucial manner, as a mistake in wording could result in misleading a patient in regards to their reports. Integration of LLMs into healthcare communications demands precise validation to avoid patient misinterpretations.

Advancing Health Systems Globally

Table 2 below outlines how NLP has the potential to advance health care systems globally by addressing and identifying health disparities among sub-groups of the population, identifying high risk patients, addressing issues of gender disparity, enhancing clinical education related to different medical conditions through better diagnostics and addressing issues that will become increasingly important in the future such as managing the patient with multi-morbid conditions.

Limitations

This narrative review has limitations that should be considered when interpreting its conclusions. Firstly, the search was limited to papers published between 2018 and 2023, which may have excluded earlier relevant papers on NLP in healthcare. Additionally, the rapid evolution of NLP technologies means that some of the reviewed studies may not reflect the most current advancements in the field. Moreover, the use of only two databases (Google Scholar and PubMed) may have limited the scope of included studies. Lastly, as this is a literature review rather than a systematic review or meta-analysis, it may not capture all relevant studies in the field.

Conclusions

While the studies discussed in this narrative review have shown the many uses of NLP and Deep Learning in public health and medicine, swift implementation of these technologies will be beneficial. The advantages of such programs must

Table 2. Summary of studies evaluating NLP potential in advancing health care systems

Lead author	Study	Relevance	Main findings
Karabacak M	From Text to Insight: A Natural Language Processing-Based Analysis of Topics and Trends in Neurosurgery	The study attempted to provide clear understanding of the research landscape in neurosurgery by using NLP techniques The use of NLP in synthesis and analysis of academic literature can aid in understanding current research.	NLP techniques were able to identify the most relevant topics as well as the least relevant topics in the current decade Use of NLP in synthesis and analysis of academic literature can aid in understanding of current research
Banerjee I	Automatic inference of BI-RADS final assessment categories from narrative mammography report findings	NLP can be used to infer BI-RADS final assessment categories (using mammogram findings in narrative form)	The classifier recognized significant contextual data that enabled the determination of the BI-RADS categorization.
Barr PJ	Precision Assessment of COVID-19 Phenotypes Using Large-Scale Clinic Visit Audio Recordings: Harnessing the Power of Patient Voice	NLP and machine learning approaches may allow rapid identification of people with a heightened chance of acquiring COVID-19. These methods may also help clinicians understand key symptoms in a healthcare setting.	Given the abundance of available data, a predictive model could be developed to enhance the accuracy of screening for those at greater risk of infection. The model could also be used to identify patient characteristics that correlate with an increase in severity.
Ryu JH	Natural Language Processing of Serum Protein Electrophoresis Reports in the Veterans Affairs Health Care System	NLP was utilized for the identification of monoclonal gammopathy from serum protein electrophoresis test data. This model's efficacy was assessed at various locations, utilizing both manual and ML-based systems.	The machine learning system achieved an AUC of 0.997. The manual system attained a 0.99 AUC. This led to the conclusion that binary classification of clinical documents may provide an opportunity for training highly accurate NLP systems.
Marafino BJ	Validation of Prediction Models for Critical Care Outcomes Using Natural Language Processing of Electronic Health Record Data	Existing models may underutilize data from the EHR for predicting outcomes in ICU patients. Accurate prediction of outcomes in such cases is crucial for clinical research. The study evaluated the impact of using NLP of clinical text data on the efficiency of a mortality prediction model.	A baseline model attained an AUC of 0.831. Augmentation with clinical measures and terms derived from NLP improved AUC to 0.922 NLP methods enhanced model performance consistently across different hospital sites ICU mortality prediction models incorporating NLP methods were found to demonstrate excellent predictive performance
Dorr DA	Prediction of Future Health Care Utilization Through Note-Extracted Psychosocial Factors	With advancing age, multimorbid individuals face elevated risks or poor healthcare outcomes. Psychosocial factors (eg, isolation, housing, financial insecurity, stress) can worsen these issues. This study was aimed at utilizing NLP to extract such factors from chart notes, and predicting the impact of these on health care utilization.	Using NLP allowed extraction of factors that enabled identification of a segment of older adults at high-risk. These techniques combined with analysis of other social determinants may allow efforts to address the negative impacts of such psychosocial factors
Jaiswal A	Identifying Risk Factors Associated With Lower Back Pain in Electronic Medical Record Free Text: Deep Learning Approach Using Clinical Note Annotations	Large number of people are afflicted by lower back pain As a result, this condition is a leading cause of disability. Developments in AI and NLP tech have created opportunities for identification of associated risk factors and management, using clinical notes. It was found that primary care clinical notes required some manipulation in order for meaningful free text analysis	Just 347 of the 2749 EHRs examined exhibited 1 or more risk factors, while the remaining did not exhibit any. Downsampling the training set improved the AUC by 2%.
Hua Y	The study was a deep learning approach toward identification of transgender and gender diverse patients from electronic health records	Accurate identification of gender identity in the EHR is essential for fair healthcare. This is especially challenging in transgender and gender diverse populations due to non availability or incomplete information. The study attempted to use NLP to create an efficient model for prediction of identity.	It was found that deep learning - integrated NLP models could be utilized for accurate prediction of gender identity when trained on EHR data This methodology is generalizable, and can be used to study other complex patient characteristics

not be understated, and despite the complications involved in their use, the need for such programs is increasing.²⁹

Widely used products from technology giants such as Meta (FB), Alphabet (Google) and Microsoft (Bing) are offering built-in NLP tools to customers. NLP powered chatbots and personal assistants are becoming increasingly common.

Much of the efficacy of NLP is based on the ability of algorithms to scan large datasets efficiently, understand human language (whether spoken or written) and mine scientific literature for insights. This is especially significant in healthcare, where huge amounts of data exists as unstructured text. Analyzing and interpreting data from such voluminous unstructured datasets offers many opportunities for identifying patterns, gaining insights and a resultant improvement in patient outcomes.

These technologies also have certain limitations that will need to be addressed before widespread adoption in healthcare settings. Technology in healthcare should be held to more stringent standards as compared to other areas. Access to high quality datasets for training the NLP models is critical to their development. Recognizing biases in data and eliminating these during training will also be crucial. Concerns about data use and privacy (HIPAA guidelines) will also need to be addressed at a policy level. Large scale adoption of NLP will require huge computational capabilities. There will also be a need to integrate these tools with existing IT systems and Artificial Intelligence (AI) models. Such limitations notwithstanding, NLP offers the promise of numerous benefits for public health officials, healthcare practitioners as well as patients. Increased focus on addressing the shortcomings will, in the long term, yield significant improvements across many realms of healthcare.

Acknowledgments

None.

Author Contribution Statements

Aadit Jerfy conceptualized the paper, wrote the first draft, and conducted the literature review. Owen Selden co-wrote the first draft and contributed to the methodology of the review. Rajesh Balkrishnan conceptualized the paper, developed the methodology and revised the first draft of manuscript.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr. Balkrishnan is currently a consultant for Otsuka Pharmaceuticals Inc. None of the other authors have any pertinent conflict of interests.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical/Consent Statement

Our study did not require an ethical board approval because it is a review.

ORCID iD

Rajesh Balkrishnan  <https://orcid.org/0000-0001-9363-9257>

References

1. Abbasgholizadeh Rahimi S, Légaré F, Sharma G, et al. Application of artificial intelligence in community-based primary health care: systematic scoping review and critical appraisal. *J Med Internet Res*. 2021;23(9):e29839. doi:10.2196/29839
2. AL-Ahdal T, Coker D, Awad H, Reda A, Żuratyński P, Khailaie S. Improving public health policy by comparing the public response during the start of COVID-19 and Monkeypox on Twitter in Germany: a mixed methods study. *Vaccines*. 2022;10(12):1985. doi:https://doi.org/10.3390/vaccines10121985
3. Amin K, Khosla P, Doshi R, Chheang S, Forman HP. Artificial intelligence to improve patient understanding of radiology reports. *Yale J Biol Med*. 2023;96(3):407-417. doi:10.59249/NKOY5498
4. Banerjee I, Bozkurt S, Alkim E, Sagreiya H, Kurian AW, Rubin DL. Automatic inference of BI-RADS final assessment categories from narrative mammography report findings. *J Biomed Inform*. 2019;92:103137. doi:10.1016/j.jbi.2019.103137
5. Barr PJ, Ryan J, Jacobson NC. Precision assessment of COVID-19 phenotypes using large-scale clinic visit audio recordings: harnessing the power of patient voice. *J Med Internet Res*. 2021;23(2):e20545. doi:10.2196/20545
6. Chae S, Davoudi A, Song J, et al. Predicting emergency department visits and hospitalizations for patients with heart failure in home healthcare using a time series risk model. *J Am Med Inform Assoc*. 2023;30(10):1622-1633. doi:10.1093/jamia/ocad129
7. Chenworth M, Perrone J, Love JS, Graves R, Hogg-Bremer W, Sarker A. Methadone and suboxone® mentions on twitter: thematic and sentiment analysis. *Clin Toxicol*. 2021;59(11):982-991. doi:10.1080/15563650.2021.1893742
8. Dorr DA, Quiñones AR, King T, Wei MY, White K, Bejan CA. Prediction of future health care utilization through note-extracted psychosocial factors. *Med Care*. 2022;60(8):570-578. doi:10.1097/MLR.0000000000001742
9. Han Y, Zhang Y, Vermund SH. Blockchain technology for electronic health records. *Int J Environ Res Public Health*. 2022;19(23):15577. doi:10.3390/ijerph192315577
10. Himmelstein G, Bates D, Zhou L. Examination of stigmatizing language in the electronic health record. *JAMA Netw Open*. 2022;5(1):e2144967. doi:10.1001/jamanetworkopen.2021.44967
11. Hua Y, Wang L, Nguyen V, et al. A deep learning approach for transgender and gender diverse patient identification in electronic health records. *J Biomed Inform*. 2023;147:104507. doi:10.1016/j.jbi.2023.104507
12. Huang Z, Tay E, Wee D, Guo H, Lim HY, Chow A. Public perception of the use of digital contact-tracing tools after the COVID-19 lockdown: sentiment analysis and opinion mining. *JMIR Form Res*. 2022;6(3):e33314. doi:10.2196/33314

13. Hu M, Conway M. Perspectives of the COVID-19 pandemic on Reddit: comparative natural language processing study of the United States, the United Kingdom, Canada, and Australia. *JMIR infodemiology*. 2022;2(2):e36941. doi:10.2196/36941
14. Hussain Z, Sheikh Z, Tahir A, et al. Artificial intelligence-enabled social media analysis for pharmacovigilance of COVID-19 vaccinations in the United Kingdom: observational study. *JMIR Public Health Surveill*. 2022;8(5):e32543. doi:10.2196/32543
15. Jaiswal A, Katz A, Nesca M, Milios E. Identifying risk factors associated with lower back pain in electronic medical record free text: deep learning approach using clinical note annotations. *JMIR Med Inform*. 2023;11:e45105. doi:10.2196/45105
16. Karabacak M, Schupper AJ, Carr MT, Hickman ZL, Margetis K. From text to insight: a natural language processing-based analysis of topics and trends in neurosurgery. *Neurosurgery*. 2023. 94(4):679-689. doi:10.1227/neu.0000000000002763
17. Keloth VK, Banda JM, Gurley M, et al. Representing and utilizing clinical textual data for real world studies: an OHDSI approach. *J Biomed Inform*. 2023;142:104343. doi:10.1016/j.jbi.2023.104343
18. Kong HJ. Managing unstructured big data in healthcare system. *Healthc Inform Res*. 2019;25(1):1-2. doi:10.4258/hir.2019.25.1.1
19. Lin SY, Cheng X, Zhang J, et al. Social media data mining of antitobacco campaign messages: machine learning analysis of facebook posts. *J Med Internet Res*. 2023;25:e42863. doi:10.2196/42863
20. Liew TM, Lee CS. Examining the utility of social media in COVID-19 vaccination: unsupervised learning of 672,133 Twitter posts. *JMIR Public Health Surveill*. 2021;7(11):e29789. doi:10.2196/29789
21. Lösch L, Zuiderent-Jerak T, Kunneman F, et al. Capturing emerging experiential knowledge for vaccination guidelines through natural language processing: proof-of-concept study. *J Med Internet Res*. 2023;25:e44461. doi:10.2196/44461
22. Marafino BJ, Park M, Davies JM, et al. Validation of prediction models for critical care outcomes using natural language processing of electronic health record data. *JAMA Netw Open*. 2018;1(8):e185097. doi:10.1001/jamanetworkopen.2018.5097
23. Niu Q, Liu J, Kato M, et al. Public opinion and sentiment before and at the beginning of COVID-19 vaccinations in Japan: Twitter analysis. *JMIR infodemiology*. 2022;2(1):e32335. doi:10.2196/32335
24. Ryu JH, Zimolzak AJ. Natural language processing of serum protein electrophoresis reports in the Veterans Affairs Health Care System. *JCO clinical cancer informatics*. 2020;4:749-756. doi:10.1200/CCI.19.00167
25. Smith TB, Vacca R, Mantegazza L, Capua I. Discovering new pathways toward integration between health and sustainable development goals with natural language processing and network science. *Global Health*. 2023;19(1):44. doi:10.1186/s12992-023-00943-8
26. Vaid A, Johnson KW, Badgeley MA, et al. Using deep-learning algorithms to simultaneously identify right and left ventricular dysfunction from the electrocardiogram. *JACC Cardiovasc Imaging*. 2022;15(3):395-410. doi:10.1016/j.jcmg.2021.08.004
27. Yang R. What is AUC? Arize AI. 2022. Accessed August 15, 2024. <https://arize.com/blog/what-is-auc/>
28. IBM. What is blockchain? n.d. Accessed March 16, 2024. <https://www.ibm.com/topics/blockchain>
29. MarketsandMarkets. NLP in healthcare & life sciences market size & forecast [latest]. 2023. <https://www.marketsandmarkets.com/Market-Reports/healthcare-lifesciences-nlp-market-131821021.html>