Artificial intelligence approaches using natural language processing to advance EHR-based clinical research in Allergy, Asthma, and Immunology

Young Juhn, MD, MPH,

Precision Population Science Lab, Division of Community Pediatric and Adolescent Medicine, Department of Pediatric and Adolescent Medicine and Division of Allergy, Department of Medicine, Mayo Clinic, Rochester, MN

Hongfang Liu, Ph.D.

Division of Digital Health, Department of Health Sciences Research, Mayo Clinic, Rochester, MN

Abstract

The wide adoption of electronic health record systems (EHRs) in health care generates big real-world data that opens new venues to conduct clinical research. As a large amount of valuable clinical information is locked in clinical narratives, natural language processing (NLP) techniques as an artificial intelligence approach have been leveraged to extract information from clinical narratives in EHRs. This capability of NLP potentially enables automated chart review for identifying patients with distinctive clinical characteristics in clinical care and reduces methodological heterogeneity in defining phenotype obscuring biological heterogeneity in research concerning allergy, asthma, and immunology. This brief review discusses the current literature on the secondary use of EHR data for clinical research concerning allergy, asthma, and immunology and highlights the potential, challenges, and implications of NLP techniques.

Keywords

Introduction

Over the past decade, electronic health records (EHRs) systems have been increasingly implemented at US hospitals and clinics. Thus, the role of artificial intelligence (AI) (computer-handled human task) as a tool for enhancing clinical care and research is becoming important in a broad range of clinical practices. For example, a few recent AI works in the field of respiratory disease and other clinical areas highlight feasibility of interpretation of pulmonary function test (PFT) results and their associated diagnosis, early identification of asymptomatic left ventricular dysfunction from 12-lead EKG, and early

Send correspondence and requests for reprints to: Young J. Juhn, M.D., M.P.H., Professor, Precision Population Science Lab, Department of Pediatric and Adolescent Medicine/Internal Medicine/Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, TEL: 507-538-1642, FAX: 507-538-1852, juhn.young@mayo.edu.

detection of atypia or carcinoma in situ from breast tissue biopsy.(1–3) These works may only reflect a small aspect of opportunities that EHR-based informatics research can offer in the future. Specifically, large amounts of detailed longitudinal patient information, including clinical history, lab tests, medications, interventions, and prognoses, are accumulated, updated, and available electronically. These large clinical databases are valuable data sources for clinical and translational research. Major initiatives have been established to exploit this crucial resource, including the Clinical and Translational Science Awards (CTSA) Program, (4) the Electronic Medical Records and Genomics (eMERGE) Network that links DNA biorepositories with EHRs,(5) the Clinical Data Research Networks (CDRN) Program supported by the Patient-Centered Outcomes Research Institute (PCORI) for comparative effectiveness research,(6) and the Observational Health Data Sciences and Informatics (OHDSI) initiative (http://ohdsi.org/), which now contains a data network with over 660 million subjects, all of which are normalized to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). The great potential of these large initiatives highly depends on efficient and effective data mining approaches leading to knowledge discovery and enhancing clinical practice,(7, 8) as a large amount of clinical information is locked in clinical narratives. For example, clinicians spend a significant amount of their time documenting clinical notes such as history of present illness, clinical history, narrative description on physical exam findings, PFT report, and radiology or operative reports. Now, an online patient portal is an important platform for communication between clinicians and patients largely in free text and a valuable source for obtaining patient reported outcomes (PRO). Experts from International Data Corporation have estimated that unstructured data accounts for more than 80 percent of currently available health care data.(9) Regarding this, natural language processing (NLP) techniques have been leveraged to extract information from clinical narratives for clinical research. Now, this capability of NLP systems enables AI-assisted cohort selection and/or matching from EHRs which may make clinical trials or studies virtual or, at least, streamlined.(10–13) For example, the 2018 National NLP Clinical Challenges (n2c2) Shared-Task and Workshop on Cohort Selection for Clinical Trials was organized for competition. One group developed a rule-based NLP system for 13 selection criteria for a clinical trial using 202 training cohort and 86 testing cohort and achieved an F-measure of 0.90 (weighted mean of a test's precision (positive predictive value] and recall [sensitivity]).(13) Given the labor-intensive nature of selecting a proper clinical trial cohort meeting the enrollment and exclusion criteria, this capability of an NLP system may potentially enable a virtual clinical trial and observational studies apart from enhancing our abilities for clinical trials. Also, as EHRs-based research as a source of real-world evidence is complimentary to traditional randomized controlled trial (RCT)-based evidence as recently recognized by FDA(14), the utility of NLP as an efficient and effective approach leveraging EHRs for real world evidence is likely to be increasingly recognized. In this review, first we will focus on discussing the current literature on NLP-based research in allergy, asthma, and immunology based on our systematic literature review. We refer readers to the recent review on broad information extraction approaches (beyond allergy, asthma, and immunology) from EHRs for further understanding.(15, 16)In the second part, we will introduce NLP to clinicians and highlight the role of NLP in data mining and knowledge discovery from EHRs using specific

case studies. We will conclude this review with discussing the implications of NLP-based clinical research and care in allergy, asthma, and immunology.

Systematic review for NLP-Based Research in Allergy, Asthma, and Immunology

To assess the current state of EHRs-based research utilizing NLP in the field of asthma, allergy, and immunology, a comprehensive literature search of several databases from January 1, 2000, to August 13th, 2019, English language, was conducted (see eTable 1 for the detailed method for our systematic literature search and the summary of results of each study). Only 21 papers were included in this systematic review by excluding abstracts, review papers, research protocols, unavailable full papers, non-EHRs or non-NLP based studies, and non-matching topics. Overall, NLP-based research in the field of asthma, allergy and immunology is limited in the literature. Herein we summarize the current state of NLP-based research leveraging EHRs for allergic disorders and discuss the key findings on the limitations of NLP-based research from the literature review.

1. Asthma:

NLP is most commonly applied for determining asthma status or outcomes as shown in eTable 1. There are a total of 13 publications which utilized NLP logic (computer follows the preset rules in the program) or NLP ML (computer learns latent rules through human guidance called annotation [supervised] or no human guidance [unsupervised]) to determine asthma status or outcomes. Apart from 9 reports from our research group,(17–25) there are only 4 publications which utilized NLP algorithms in asthma status or outcomes(26–28). As shown in eTable 1, overall, performance of NLP systems for asthma status and outcomes appear to vary depending on the nature of asthma concepts, outcomes, and NLP methods. Our prior work primarily applied NLP (logic or ML) to detect asthma status or prognosis based on predetermined criteria instead of searching free text terms as discussed in the following section.

2. Atopic dermatitis:

Only one paper utilized NLP to define atopic dermatitis based on a predetermined criteria(29). In this study, the authors developed a machine learning algorithm for identifying atopic dermatitis in adults from EHRs using Hanifin-Rajka and UK Working Party criteria. The study used three non-overlapping data subsets stratified by class label: training (60%), validation (20%), testing (20%) of 562 adult patients. They reported precision (69–84%), recall (51–75%) and F (0.62–0.79) depending on structured code only, NLP only, and structured code-NLP and type of atopic dermatitis criteria.

3. Allergy entry in EHRs:

Another area to which NLP has been applied is identification of allergy entry in EHRs. Even if standard terminologies for allergy entry could be searched and retrieved, such terminologies exist in an inconsistent or non-standardized format (ie, lexical variation) posing challenges for identifying specific allergy status from EHRs For example,

Systematized Nomenclature of Medicine—Clinical Terms (SNOMED-CT) (standardized clinical terminology system developed and maintained by SNOMED International, a non-profit international organization for the purpose of information entry into EHRs [ICD is a classification system for information retrieval]) was used to match terms referring to allergy entry in EHRs but matching tends to be low (e.g., one study reported 82% for SNOWMED-CT) because 6% of all allergies and 48% of food allergies were in a free-text format.(30) Because of this reason, NLP had to be used for addressing this challenge (2nd most common use cases for NLP after asthma), detection of specific allergy status from EHRs. There are 4 NLP-based studies (3 studies based on NLP-logic for food allergy,(30) any allergy entries in ED setting(31) and food/drug allergy(32) and 1 study based on supervised NLP-ML for any allergy misspelling(33)). Their overall performance of NLP algorithms was excellent (F-measures [mean of positive predictive value and sensitivity] of 94–99%(30, 32, 33) or slightly lower F-measures of 80–92% depending on allergies(31)). These are promising results for application to clinical practice.

4. Other areas:

There is one paper for detecting smoking status related to asthma exacerbation in EHRs based on NLP-logic with positive predictive value of 79% and specificity of 90%.(34) One paper used NLP and multiple ML classifiers to extract and identify anaphylaxis incidence from unstructured data (F-measure: 048–0.95)(35). Another paper used NLP to extract and determine the incidence of hypersensitivity reaction to non-steroidal anti-inflammatory drug (NSAID) from EHRs.(36)

There are several key findings for the current state of NLP-based research in the field of asthma, allergy, and immunology.

First, overall, NLP is under-utilized in the field of asthma, allergy, and immunology. While there is some NLP-based research in allergy and asthma, little research based on NLP approaches has been done in atopic dermatitis or allergic rhinitis. Apart from NLP algorithms for asthma criteria and prognosis reported by our group discussed later,(20, 22, 23, 37, 38) NLP algorithms for complex concepts such as predetermined criteria for allergic disorders are significantly limited recognizing that valuable information required for complex predetermined criteria for allergic disorders exists in a free-text format of EHRs. NLP is a useful tool for identifying a cohort with distinctive clinical characteristics in clinical conditions with significant heterogeneity such as asthma. For example, as NLP has a technical feature that researchers may modify the existing NLP algorithms for a condition (eg, asthma) to identify a subgroup of patients with distinctive clinical characteristics (eg, atopic asthma or asthmatics with impaired lung function), which may not be feasible with ICD codes for asthma.

Second, the current literature on EHR-based research using NLP has a tendency to not capture and utilize a temporal component for outcome events of interest. As NLP leverages free texts of clinician's notes in EHRs, it is more accurate for detecting index date of allergic disorders than structured code. As the status of allergic disorders in large-scale clinical studies is frequently determined by a survey question about a physician diagnosis of asthma (eg, "have you ever been told that you (or your child) have a diagnosis of asthma?") without

knowing the index date, NLP is useful for detecting index date (when one fulfilled predetermined criteria as shown in our research work discussed later) of asthma or other outcomes such as remission on a large scale. For clinical care and research, it is crucially important to determine the index or incidence date of outcome events (e.g., asthma onset, asthma remission or relapse, etc.) as temporality is an essential aspect to discern the relationship between predictors and outcomes. In this respect, NLP systems have the capability for capturing temporal components of outcome events over time as NLP programs detect temporal components for each note for a specific concept.

Third, as one of the major challenges in informatics research, portability and external validation of AI algorithms are often difficult to assess due to patient confidentiality and technology transfer. Only a few studies have addressed this challenge. Clinical practice and workflow vary across institutions, which results in different practice settings and reporting schemes for generating EMRs. Also, it has been demonstrated that clinical language is not homogeneous, but instead consists of heterogeneous characteristics (eg, syntactic variation or semantic variation).(39) All these factors pose challenges for applying an NLP algorithm developed from one institution to another and need to be considered in portability of NLP algorithms. This is an active research area for addressing these challenges.

Finally, not least, apart from its accuracy of EHR-based AI tools, another important aspect for assessing properness or utility of such tools is acceptability, user interface, usability, or workflow for the entire care process to ensure EHR-based AI tools are helpful rather than a task or burden to clinicians or researchers. Dr. Homer Warner, a pioneer in medical informatics said "Medical informatics = 10% technology + 10% medicine + 80% sociology" highlighting the crucial importance of designing and implementing informatics tools including NLP considering patient-centered, clinician-empowering, and team-based workflow. In this respect, there is no widely accepted standard metrics assessing or reporting these important domains. We further discussed this important issue and challenge in the later section.

The Case for Natural Language Processing Systems as an EHR-Based Clinical Research Tool

1. Introduction for NLP:

We discuss a brief overview of NLP to introduce the conceptual understanding of NLP for clinicians but do not intend to cover the topic of NLP for the purpose of performing NLP-based research. We refer readers to a few review papers that discuss NLP in depth.(40–42) NLP is a field in computer science, artificial intelligence, and computational linguistics which enables the interactions between computers and human languages and bridges the gap between clinical human language and computational systems. NLP, machine learning and deep learning (a field of machine learning) all are part of the field of AI, and NLP utilizes machine learning and deep learning. NLP is computer-handled communication with humans through natural languages (human languages, not computer languages). NLP can be broadly defined as any tasks that computationally utilize human languages such as written language (texts) or spoken language (speech) to detect the underlying concepts. In the context of

clinical research, we limit it to EHRs in the clinical setting, which only encompass protected health information of patients in digital format created by patients, clinicians, and a broad range of other care team members. There are two approaches for NLP: one is being operated by rule-based (computer follows the preset rules in the program) vs. machine learning (ML) approaches (computer learns latent rules through human guidance called annotation [supervised] or no human guidance [unsupervised]). Experts from International Data Corporation have estimated that unstructured data accounts for more than 80 percent of currently available health care data.(9) While valuable clinical information (e.g., history of present illness, clinical history, narrative description on physical exam findings, PFT report, and radiology or operation reports) is embedded in free texts of EHRs, it is labor intensive to extract such information and enter them into structured data. Thus, most, if not all, large-scale studies or population management strategies heavily rely on structured data and do not tap into unstructured data in EHRs in a way preventing clinicians and researchers from complete and nuanced information in EHRs. There has been a concern about underutilization of unstructured data in EHRs leading to persistent use of inaccurate structured data (e.g., ICD codes) in both clinical research and care.(37, 43) NLP can potentially address this challenge.

As shown in Figure 1, conceptually, NLP algorithms for cohort identification *extract* the information (or concepts) from EHRs, then *process* the extracted information, and finally *classify* patients into a subgroup according to rules or learners. These conceptual procedures for NLP are rather complex because NLP is not a single technique; rather, it consists of multiple techniques grouped together. To map words or phrases to concepts of interest, it requires careful text pre-processing and other NLP tasks to convert a document from its natural form into a bag of words. (40, 41) A few examples of low-level NLP tasks (text pre-processing) include: 1. sentence boundary detection which is typically defined by a period (title such as "Dr." might complicates this task), 2. tokenization (breaking a sentence into individual tokens) (eg, converting a sentence, "patients have multiple history of wheezing episodes" into "patients", "have", "multiple", "history", "wheezing", "episodes"). 3. Stemming (reducing word into a root form) (eg, converting "starting, started, and starts" to "start") and lemmatization (process mapping a token) (eg, mapping "RBC, red blood cells" to "erythrocyte"). A few examples of higher-level NLP tasks include: 1. Named entity recognition (NER, identifying specific words or phrases such as disease, genes, or medication), 2. Setting up negation rules (eg, no history, denies history, absence of history). NLP is different from a simple key word search as a simple key word search will not be able to distinguish various forms of clinician's notes in EHRs. For example, "patient has a known history of asthma" vs. " there is no history of asthma", "patient reports a wheezing episode" vs. "patient denies a wheezing episode" and "patient has a history of asthma" vs. "sister has a history of asthma but no history of asthma in the patient". To differentiate these two sentences, NLP program applies a list of negation rules in addition to other NLP tasks described above. The National Library of Medicine provides several well-known knowledge infrastructure resources that facilitate these NLP tasks (eg, UMLS Metathesaurus for records synonyms and NER, and a text collection for word disambiguation).(41) Once information extraction (IE) and processing information from EHRs are completed, which occurs at a document level, aggregation and classification take place at a patient level based on pre-

established rules (eg, asthma criteria) (ie, a part of process developing NLP algorithm). A common gold standard being used for NLP algorithm development is expert-based classification by annotating each specific feature (eg, wheezing episodes or physician diagnosis of asthma) by experts. This step typically requires some kind of annotation tool (eg, Anafora). Typically, training computer or learner takes multiple iterative process until computer correctly identifies (extract, process, and classify) the concept of interest using expert's annotation as a reference. NLP algorithms are developed or trained using a *developing or training cohort.* Once it reaches optimal performance in detecting the concept, NLP algorithm is tested using an independent *testing cohort* (eg, a k-fold cross-validation method [e.g., train/test split]). After the initial development and validation of NLP algorithms, it is necessary for investigators to further assess portability and external validation of NLP algorithms which is one of the most challenging aspects before its implementation.

2. Use cases for NLP system in asthma research:

As discussed earlier, as an example, we recently developed and validated NLP algorithms for two existing retrospective criteria for asthma: Predetermined Asthma Criteria (PAC) and Asthma Predictive Index (API) described in Table 1 (20, 22, 23, 37) and asthma prognosis (. (38) Given the significant heterogeneity in determining asthma status for asthma care and research (eg, 60 different definitions in the literature) and the limited research leveraging free texts in EHRs for asthma research, our group developed NLP algorithms for PAC and API by conducting a retrospective birth cohort study (training cohort=430 and testing cohort=500 from the 1997–2007 Mayo Clinic Birth Cohort[n=8,525]) which used clinician's manual chart review with annotation to apply both criteria as gold standard and asthma status by NLP algorithms as a predictor. Despite the complex concepts of NLP-PAC to be captured from EHRs, NLP-PAC showed that sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for asthma status by clinician's manual reviews were 97%, 95%, 90%, and 98%, respectively. Therefore, our prior work on NLP algorithms for multiple asthma criteria demonstrates feasibility for developing NLP algorithms for patient-level classification of asthma status by applying complex clinical concepts on large-scales. Furthermore, we demonstrated portability (generalizability) of NLP algorithms developed at our institution across different practice settings (different geographic area, patient population, documentation practice, and EHR system) by conducting a similar birth cohort study (training cohort=298 and testing cohort=297) in Sioux Falls, South Dakota.(20, 39) Sensitivity, specificity, PPV, and NPV for the NLP-PAC algorithm in predicting asthma status by manual chart review were 92%, 96%, 89%, and 97%, respectively in a different practice setting. (20) Importantly, NLP-PAC and NLP-API identify a group of asthmatic children with distinctive clinical and immunological characteristics (eg, persistent asthma, increased risk of asthma exacerbation, impaired lung function, increased risk of infection and Th2-high immune profile among asthmatic children who met both NLP-PAC and NLP-API).(44)

For NLP algorithm for asthma prognosis (long-term remission, relapse, and persistent asthma; remission was defined by the absence of asthma events for 3 years consecutively) after the index date of asthma, sensitivity, PPV and F-measure for asthma events were 94%,

98%, and 0.96%, respectively. An important feature of our NLP system for asthma criteria and prognosis is capability of detecting index date of asthma or outcome events (i.e., onset of asthma or remission or relapse) which is crucially important for an epidemiological investigation requiring temporality. These capabilities of NLP algorithms may enable scalable precision care or population management and reduce methodological heterogeneity in phenotyping in allergy, asthma, and immunology(45) which often obscures true biological heterogeneity deterring translation of scientific findings into practice while reducing the burden of clinicians and researchers.

Implications of NLP for EHR-Based on Clinical Research and Care

Figure 2 illustrates modular components of EHRs (top row) which encompass specific clinical information in both structured (eg, lab data) and non-structured (eg, clinical notes) formats (bottom row) and the multimodal capability of NLP can extract and process this information, and classify patients enabling important clinical tasks (right column) such as quality report and clinical decision support. Thus, NLP might be an important method to make EHRs a helpful data source for addressing the needs and supporting important activities of clinicians and researchers while reducing their burden for chart review and data mining from EHRs (e.g., 70% of clinicians using EHRs reported burnout). (7, 8) For example, clinicians often need to meet regulatory compliance (eg, reporting asthma care quality measure to a state agency, Minnesota Community Measure(46)) which requires accurate determination of children with asthma as a denominator for the measure. NLP for asthma criteria could help clinicians efficiently complete this task while they still have control over asthma status suggested by NLP as NLP algorithms provide clinicians with evidence (ie, clinical notes supporting asthma diagnosis) for its logic. Another example is NLP might be able to retrieve and summarize pertinent information from EHRs for clinicians enabling automated chart review leveraging free texts in EHRs, and thus, clinicians can make their management decision for asthma efficiently by not relying on manual chart review. Our recent study showed automated chart review via NLP system reduced clinician's time for making their clinical decision by about 80%(47). Application of AI or other IT technology to clinical care and research is likely to be an important direction in the US health care system (e.g., 21 Century Cures Act in 2016).(48) As performance of AI algorithms largely depends on high-quality input data, and thus, NLP as a data mining tool for EHRs will be a cornerstone for achieving high-value care and high-quality research reducing methodological heterogeneity in phenotyping, a key step for advancing allergy and asthma research. The current interest and effort for establishing the interoperability standards for EHRs led by the Office of the National Coordinator for Health Information Technology under the 21 Century Cures Act has important implications on future EHR-driven clinical care and research. (49, 50)

At the same time, there are also many challenges for NLP-based research to be addressed in the future: data quality issues (eg, missing, biases, etc.), privacy issue hosting or transferring EHR data to cloud, algorithmic bias, lack of interoperability standards, and health information technology (HIT)-related clinician burnout and workflow issues. For example, the 2019 NIH workshop on Machine Intelligence in Health Care highlighted trustworthiness, explainability, usability, transparency, and fairness of AI as potential challenges(51). While

large-scale studies for asthma or allergy are primarily based on structured data (eg, ICD code) and are being routinely performed despite the known limitations, as discussed in our systematic review for the current literature, at present, research work leveraging free texts in EHRs via NLP is severely limited. While NLP is not free from systematic biases from EHRs (52, 53), it may help us to identify algorithmic biases stemming from systematic biases of EHRs. Despite such limitations, EHRs-based research as a source of real-world evidence is complimentary to traditional RCT-based evidence as recently recognized by FDA.(14) Dr. Edward Shortliffe recently provided useful guidance for the-must-have-features of an informatics tool if it is to be accepted and integrated into routine workflow(54): 1. Black boxes are unacceptable (should provide why and how), 2. Time is a scarce resource, 3. Complexity and lack of usability thwart use, 4. Relevance and insight are essential, and 5. Delivery of knowledge and information must be respectful.

Given the current trends, impact and potential of EHR-based research, innovative AI approaches for data mining, and knowledge discovery from EHRs such as NLP should be pursued in education and training, clinical practice, and research in the field of allergy, asthma, and immunology. Recognizing unintended consequences, HIT needs to be carefully and properly designed and implemented in a greater socioecological context of clinical care settings than a specific HIT innovation itself.(55)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

Abbreviations

EHRs	Electronic health records
NLP	Natural language processing
AI	Artificial intelligence
ML	Machine learning
PFT	Pulmonary function test
HIT	Health information technology
PAC	Predetermined asthma criteria
API	Asthma predictive index

CTSA	Clinical and Translational Science Awards
eMERGE	Electronic Medical Records and Genomics Network
CDRN	Clinical Data Research Networks
PCORI	Patient-Centered Outcomes Research Institute
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
CDM	Common Data Model
SNOMED-CT	Systematized Nomenclature of Medicine—Clinical Terms
PPV	Positive Predictive Value
NPV	Negative Predictive Value

References

1. Topalovic M, Das N, Burgel PR, Daenen M, Derom E, Haenebalcke C, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. Eur Respir J. 2019;53(4).

2. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med. 2019;25(1):70–4. [PubMed: 30617318]

3. Mercan E, Mehta S, Bartlett J, Shapiro LG, Weaver DL, Elmore JG. Assessment of Machine Learning of Breast Pathology Structures for Automated Differentiation of Breast Cancer and High-Risk Proliferative Lesions. JAMA Network Open. 2019;2(8):e198777–e. [PubMed: 31397859]

4. NF S Clinical and Translational Science Awards. National Center for Research Resources 2011.

5. Lemke AA, Wu JT, Waudby C, Pulley J, Somkin CP, Trinidad SB. Community engagement in biobanking: Experiences from the eMERGE Network. Genomics Soc Policy. 2010;6(3):35–52.

6. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc. 2014;21(4):578–82. [PubMed: 24821743]

7. Gardner RL, Cooper E, Haskell J, Harris DA, Poplau S, Kroth PJ, et al. Physician stress and burnout: the impact of health information technology. J Am Med Inform Assoc. 2019;26(2):106–14. [PubMed: 30517663]

8. Economics M. EHRs shoudl be a tool not a task 2019 [Available from: https://www.medicaleconomics.com/article/ehrs-should-be-tool-not-task.

9. Martin-Sanchez F, Verspoor K. Big data in medicine is driving big changes. Yearb Med Inform. 2014;9(1):14–20. [PubMed: 25123716]

10. Challenge. NCNNC. National NLP Clinical Challenge 2019 [Available from: N2C2: National NLP Clinical Challenges https://n2c2.dbmi.hms.harvard.edu/ Accessed September 23, 2019.

11. Xiong Y, Shi X, Chen S, Jiang D, Tang B, Wang X, et al. Cohort selection for clinical trials using hierarchical neural network. Journal of the American Medical Informatics Association. 2019.

12. Fang G, Annis IE, Elston-Lafata J, Cykert S. Applying machine learning to predict real-world individual treatment effects: insights from a virtual patient cohort. Journal of the American Medical Informatics Association. 2019;26(10):977–88. [PubMed: 31220274]

13. Chen L, Gu Y, Ji X, Lou C, Sun Z, Li H, et al. Clinical trial cohort selection based on multi-level rule-based natural language processing system. Journal of the American Medical Informatics Association. 2019;26(11):1218–26. [PubMed: 31300825]

14. Jarow JP, LaVange L, Woodcock J. Multidimensional evidence generation and fda regulatory decision making: Defining and using "real-world" data. Jama. 2017.

15. Wang Y, Wang L, Rastegar-Mojarad M, Moon S, Shen F, Afzal N, et al. Clinical information extraction applications: A literature review. Journal of Biomedical Informatics. 2018;77:34–49. [PubMed: 29162496]

16. Shivade C, Raghavan P, Fosler-Lussier E, Embi PJ, Elhadad N, Johnson SB, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc. 2014;21(2):221–30. [PubMed: 24201027]

17. Wu ST, Juhn YJ, Sohn S, Liu H. Patient-level temporal aggregation for text-based asthma status ascertainment. J Am Med Inform Assoc. 2014;21(5):876–84. [PubMed: 24833775]

18. Wu ST, Sohn S, Ravikumar KE, Wagholikar K, Jonnalagadda SR, Liu H, et al. Automated chart review for asthma cohort identification using natural language processing: an exploratory study. Ann Allergy Asthma Immunol. 2013;111(5):364–9. [PubMed: 24125142]

19. Wu S, Liu S, Sohn S, Moon S, Wi C-i, Juhn Y, et al. Modeling asynchronous event sequences with RNNs. Journal of Biomedical Informatics. 2018;83:167–77. [PubMed: 29883623]

20. Wi CI, Sohn S, Ali M, Krusemark E, Ryu E, Liu H, et al. Natural Language Processing for Asthma Ascertainment in Different Practice Settings. J Allergy Clin Immunol Pract. 2018;6(1):126–31. [PubMed: 28634104]

21. Sohn S, Wi C-I, Wu ST, Liu H, Ryu E, Krusemark E, et al. Ascertainment of asthma prognosis using natural language processing from electronic medical records. Journal of Allergy and Clinical Immunology. 2018;141(6):2292–4.e3.

22. Kaur H, Sohn S, Wi CI, Ryu E, Park MA, Bachman K, et al. Automated chart review utilizing natural language processing algorithm for asthma predictive index. BMC Pulm Med. 2018;18(1):34. [PubMed: 29439692]

23. Wi CI, Sohn S, Rolfes MC, Seabright A, Ryu E, Voge G, et al. Application of a Natural Language Processing Algorithm to Asthma Ascertainment. An Automated Chart Review. Am J Respir Crit Care Med. 2017;196(4):430–7. [PubMed: 28375665]

24. Sohn S, Wi CI, Juhn YJ, Liu H. Analysis of Clinical Variations in Asthma Care Documented in Electronic Health Records Between Staff and Resident Physicians. Stud Health Technol Inform. 2017;245:1170–4. [PubMed: 29295287]

25. Sohn S, Wang Y, Wi C-I, Krusemark EA, Ryu E, Ali MH, et al. Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. Journal of the American Medical Informatics Association. 2017:ocx138–ocx.

26. Afzal Z, Engelkes M, Verhamme KM, Janssens HM, Sturkenboom MC, Kors JA, et al. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases. Pharmacoepidemiol Drug Saf. 2013;22(8):826–33. [PubMed: 23592573]

27. Himes BE, Kohane IS, Ramoni MF, Weiss ST. Characterization of patients who suffer asthma exacerbations using data extracted from electronic medical records. AMIA Annu Symp Proc. 2008:308–12. [PubMed: 18999057]

28. Liang H, Tsui BY, Ni H, Valentim CCS, Baxter SL, Liu G, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nature Medicine. 2019;25(3):433–8.

29. Gustafson E, Pacheco J, Wehbe F, Silverberg J, Thompson W. A Machine Learning Algorithm for Identifying Atopic Dermatitis in Adults from Electronic Health Records. IEEE Int Conf Healthc Inform 2017;2017:83–90.

30. Plasek JM, Goss FR, Lai KH, Lau JJ, Seger DL, Blumenthal KG, et al. Food entries in a large allergy data repository. J Am Med Inform Assoc. 2016;23(e1):e79–87. [PubMed: 26384406]

31. Goss FR, Plasek JM, Lau JJ, Seger DL, Chang FY, Zhou L. An evaluation of a natural language processing tool for identifying and encoding allergy information in emergency department clinical notes. AMIA Annu Symp Proc. 2014;2014:580–8. [PubMed: 25954363]

32. Epstein RH, St Jacques P, Stockin M, Rothman B, Ehrenfeld JM, Denny JC. Automated identification of drug and food allergies entered using non-standard terminology. J Am Med Inform Assoc. 2013;20(5):962–8. [PubMed: 23748627]

33. Lai KH, Topaz M, Goss FR, Zhou L. Automated misspelling detection and correction in clinical free-text records. J Biomed Inform. 2015;55:188–95. [PubMed: 25917057]

34. Meystre SM, Deshmukh VG, Mitchell J. A clinical use case to evaluate the i2b2 Hive: predicting asthma exacerbations. AMIA Annu Symp Proc. 2009;2009:442–6. [PubMed: 20351896]

35. Segura-Bedmar I, Colon-Ruiz C, Tejedor-Alonso MA, Moro-Moro M. Predicting of anaphylaxis in big data EMR by exploring machine learning approaches. J Biomed Inform. 2018;87:50–9. [PubMed: 30266231]

36. Blumenthal KG, Lai KH, Huang M, Wallace ZS, Wickner PG, Zhou L. Adverse and Hypersensitivity Reactions to Prescription Nonsteroidal Anti-Inflammatory Agents in a Large Health Care System. J Allergy Clin Immunol Pract. 2017;5(3):737–43.e3. [PubMed: 28110055]

37. Seol HYSS, Liu H, Wi C, Ryu E, Park MA, Juhn YJ. Early Identification of Childhood Asthma: The Role of Informatics in an Era of Electronic Health Records Frontiers in Pediatrics. 2019

38. Sohn S, Wi CI, Wu ST, Liu H, Ryu E, Krusemark E, et al. Ascertainment of asthma prognosis using natural language processing from electronic medical records. J Allergy Clin Immunol 2018 2 10 [Epub ahead of print]. 2018.

39. Sohn S, Wang Y, Wi CI, Krusemark EA, Ryu E, Ali MH, et al. Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. J Am Med Inform Assoc. 2017.

40. Kimia AA, Savova G, Landschaft A, Harper MB. An Introduction to Natural Language Processing: How You Can Get More From Those Electronic Notes You Are Generating. Pediatr Emerg Care. 2015;31(7):536–41. [PubMed: 26148107]

41. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc. 2011;18(5):544–51. [PubMed: 21846786]

42. Friedman C, Rindflesch TC, Corn M. Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the National Library of Medicine. Journal of Biomedical Informatics. 2013;46(5):765–73. [PubMed: 23810857]

43. Pak Hon S.. Unstructured data in health care, : Healthcare Tech; 2019 [Available from: https://artificial-intelligence.healthcaretechoutlook.com/cxoinsights/unstructured-data-in-healthcare-nid-506.html.

44. Seol HY, MC R, Wi C, Sohn S, Ryu E, Park MA, et al. Expert Artificial Intelligence-based Natural Language Processing Characterizes Childhood Asthma BMJ Open Respiratory Research (In Press). 2020 10.1136/bmjresp-2019-000524

45. Van Wonderen KE, Van Der Mark LB, Mohrs J, Bindels PJE, Van Aalderen WMC, Ter Riet G. Different definitions in childhood asthma: how dependable is the dependent variable? European Respiratory Journal. 2010;36(1):48.

46. Minnesota Community Measurement. 2017 Health Care Quality Report. Minnesota Community Measurement; 2017.

47. Juhn YWC, Sohn S, Ryu E, Park M, Fladager Muth J, Seol HY, Moon S, King K, Wheeler P, Liu H, Ihrke K, McWilliams D,. Asthma-Guidance and Prediction System (a-GPS) As a Precision Asthma Care Tool Annual American Academy of Allergy, Asthma and Immunology; Philadelphia, PA2020.

48. (HIMSS) HIaMSS. 21st Century Cures Act - A Summary: HIMSS; 2017 [Available from: https://www.himss.org/news/21st-century-cures-act-summary.

49. Tahir Darius. Provider Groups: Slow down info blocking rules: Politico; 2019 [Available from: https://www.politico.com/newsletters/morning-ehealth/2019/09/18/provider-groups-slow-down-info-blocking-rules-478847.

50. Cohen Jessica. ONC working on app privacy with Congress and White House 2019 [Available from: https://www.modernhealthcare.com/politics-policy/rucker-onc-working-app-privacy-congress-white-house.

51. Bioengineering NIoBIa. Machine Intelligence in Healthcare: NIH; 2019 [Available from: https://ncats.nih.gov/expertise/machine-intelligence#workshop.

52. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447. [PubMed: 31649194]

53. Crown WH. Potential application of machine learning in health outcomes research and some statistical cautions. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2015;18(2):137–40. [PubMed: 25773546]

54. Shortliffe EH, Sepúlveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. Jama. 2018;320(21):2199–200. [PubMed: 30398550]

55. Bakken S Building the evidence base on health information technology–related clinician burnout: a response to impact of health information technology on burnout remains unknown—for now. Journal of the American Medical Informatics Association. 2019.
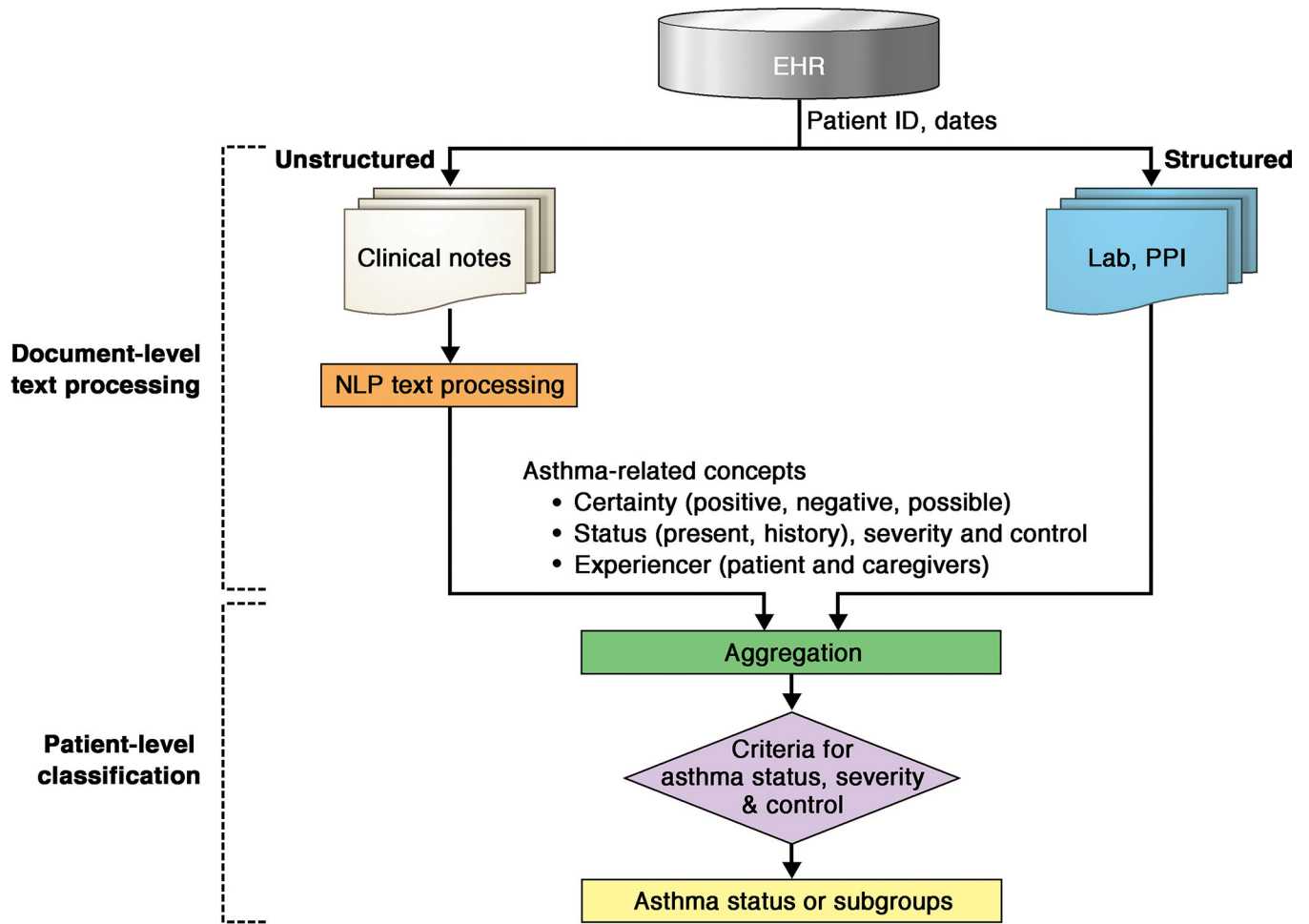
FIG 1.
NLP Algorithms for document-level text processing and patient-level classification. NLP algorithms for cohort identification extract the information (or concepts) from EHRs, then process the extracted information, and finally classify patients into a subgroup according to rules or learners. Once information extraction and information processing from EHRs are completed, which occurs at a document level, aggregation and classification take place at a patient level on the basis of preestablished rules (eg, asthma criteria) (ie, a part of process developing NLP algorithm). PPI, Patient-provided information
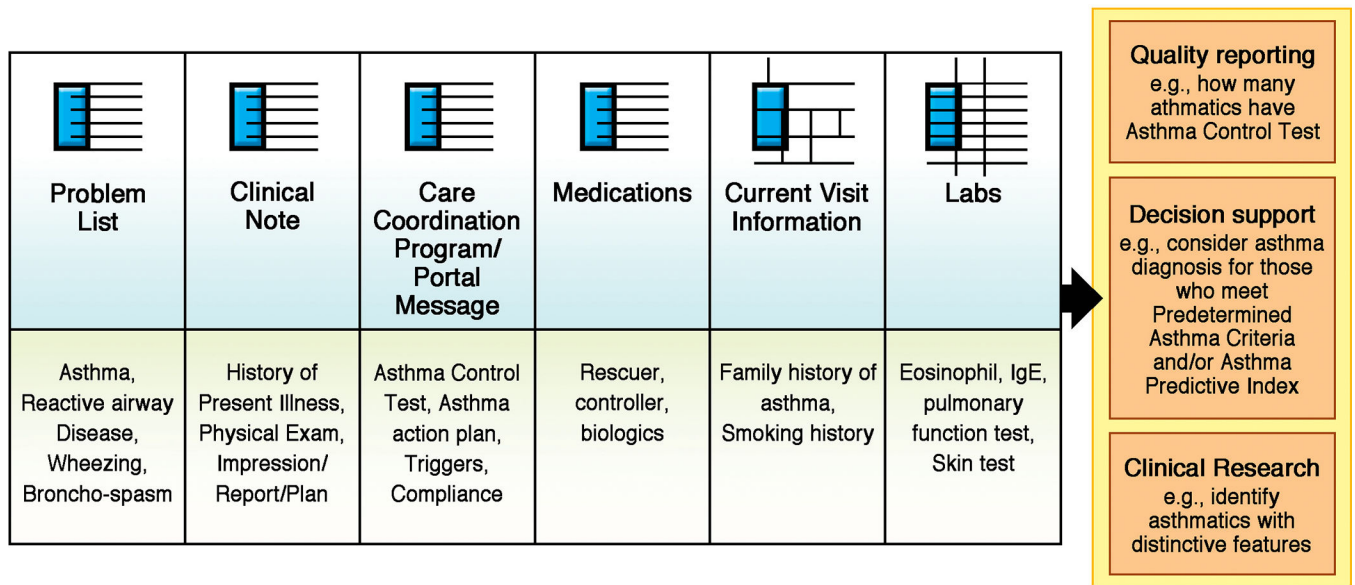
Problem List	Clinical Note	Care Coordination Program/ Portal Message	Medications	Current Visit Information	Labs
Asthma, Reactive airway Disease, Wheezing, Broncho-spasm	History of Present Illness, Physical Exam, Impression/ Report/Plan	Asthma Control Test, Asthma action plan, Triggers, Compliance	Rescuer, controller, biologics	Family history of asthma, Smoking history	Eosinophil, IgE, pulmonary function test, Skin test

Quality reporting
e.g., how many athmatics have Asthma Control Test

Decision support
e.g., consider asthma diagnosis for those who meet Predetermined Asthma Criteria and/or Asthma Predictive Index

Clinical Research
e.g., identify asthmatics with distinctive features

FIG 2.

NLP Use cases for asthma. The figure illustrates modular components of EHRs, which encompass specific clinical information in both structured (eg, laboratory data) and nonstructured (eg, clinical notes) formats. It also shows the multimodal capability of NLP, which can extract and process this information, and classify patients, enabling important clinical tasks such as quality report and clinical decision support.

TABLE I.

Two asthma criteria

A. Predetermined asthma criteria (PAC)

Patients were considered to have *definite* asthma if a physician had made a diagnosis of asthma and/or if each of the following three conditions were present, and they were considered to have *probable* asthma if only the first two conditions were present:

1. History of cough with wheezing, and/or dyspnea, OR history of cough and/or dyspnea plus wheezing on examination,

2. Substantial variability in symptoms from time to time or periods of weeks or more when symptoms were absent, and

3. Two or more of the following:

• Sleep disturbance by nocturnal cough and wheeze

• Nonsmoker (14 years or older)

• Nasal polyps

• Blood eosinophilia higher than 300/uL

• Positive wheal and flare skin tests OR elevated serum IgE

•History of hay fever or infantile eczema OR cough, dyspnea, and wheezing regularly on exposure to an antigen

• Pulmonary function tests showing one FEV_1 or FVC less than 70% predicted and another with at least 20% improvement to an FEV_1 of higher

than70% predicted OR methacholine challenge test showing 20% or greater decrease in FEV_1[1]

• Favorable clinical response to bronchodilator

B. Asthma Predictive Index (API)

Major Criteria	Minor Criteria
1. Physician diagnosis of asthma for parents	1. Physician diagnosis of allergic rhinitis for patient
2. Physician diagnosis of eczema for patient	2. Wheezing apart from colds
	3. Eosinophilia (4%)

FVC, forced vital capacity; and FEV_1, forced expiratory volume in 1 sec. API (+): Frequent wheezing episodes (e.g., two or more wheezing episodes per year) plus at least one of two major criteria or two of three minor criteria