

Original Paper

# Implementation of Machine Learning Pipelines for Clinical Practice: Development and Validation Study

Lara J Kanbar<sup>1</sup>, PhD; Benjamin Wissel<sup>2</sup>, PhD; Yizhao Ni<sup>2,3</sup>, PhD; Nathan Pajor<sup>1,2,3</sup>, MD; Tracy Glauser<sup>3,4</sup>, MD; John Pestian<sup>2,3</sup>, PhD; Judith W Dexheimer<sup>2,3,5</sup>, PhD

<sup>1</sup>Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

<sup>2</sup>Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

<sup>3</sup>Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States

<sup>4</sup>Division of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

<sup>5</sup>Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

**Corresponding Author:**

Judith W Dexheimer, PhD

Division of Biomedical Informatics

Cincinnati Children's Hospital Medical Center

3333 Burnet Avenue

Cincinnati, OH, 45229

United States

Phone: 1 5138032962

Email: [judith.dexheimer@cchmc.org](mailto:judith.dexheimer@cchmc.org)

## Abstract

**Background:** Artificial intelligence (AI) technologies, such as machine learning and natural language processing, have the potential to provide new insights into complex health data. Although powerful, these algorithms rarely move from experimental studies to direct clinical care implementation.

**Objective:** We aimed to describe the key components for successful development and integration of two AI technology-based research pipelines for clinical practice.

**Methods:** We summarized the approach, results, and key learnings from the implementation of the following two systems implemented at a large, tertiary care children's hospital: (1) epilepsy surgical candidate identification (or epilepsy ID) in an ambulatory neurology clinic; and (2) an automated clinical trial eligibility screener (ACTES) for the real-time identification of patients for research studies in a pediatric emergency department.

**Results:** The epilepsy ID system performed as well as board-certified neurologists in identifying surgical candidates (with a sensitivity of 71% and positive predictive value of 77%). The ACTES system decreased coordinator screening time by 12.9%. The success of each project was largely dependent upon the collaboration between machine learning experts, research and operational information technology professionals, longitudinal support from clinical providers, and institutional leadership.

**Conclusions:** These projects showcase novel interactions between machine learning recommendations and providers during clinical care. Our deployment provides seamless, real-time integration of AI technology to provide decision support and improve patient care.

(*JMIR Med Inform* 2022;10(12):e37833) doi: [10.2196/37833](https://doi.org/10.2196/37833)

**KEYWORDS**

electronic health record; natural language processing; epilepsy; clinical decision support; machine learning; emergency medicine; artificial intelligence

## Introduction

With the rampant growth in health data, artificial intelligence (AI) technologies, such as machine learning and natural

language processing (NLP), provide a powerful means to extract meaningful associations from big data sets [1]. Applications of machine learning are far-reaching and have included patient identification, computer vision, speech recognition, web searches, and phenotype discovery [2-9].

The electronic health record (EHR) captures data relating to clinical encounters, but as much as 30%-50% of these data are available only in free text [10]. As such, one particularly valuable means to understand health care data involves NLP. NLP is a technique of incorporating free-text analysis and statistical methods into computerized algorithms to derive linguistic features (eg, physicians' diagnosis) from human language input [11]. Clinical care and research can benefit from using this unstructured text information [12,13]. NLP has been used for surveillance, adverse event detection [14-18], medication identification [19], and extraction of data from radiology reports [20-22]. NLP has also successfully been applied to evaluate clinical notes and provide recommendations as part of clinical decision support (CDS) tools [23].

These CDS tools can change user behavior; however, to ensure successful implementation, user involvement in CDS design is critical [24-30]. CDS tools using AI and NLP technologies remain less implementable directly into real-time clinical care with long-term success [31-34]. The reason integration of these AI pipelines within a clinical health system is challenging is that it requires coordination with the following: (1) key stakeholders and expected end users of the CDS tools; (2) biomedical informatics professionals who design the AI; (3) research information technology (IT) professionals who design the CDS tools with stakeholders in mind; and (4) operational IT professionals who are responsible for maintenance, uptime, and EHR integration [35].

In this work, we report the main modifications implemented to improve the development and real-time integration of two AI technology-based pipelines using NLP in a tertiary pediatric health care institution. These modifications contributed to the successful deployment and ongoing utilization of these pipelines.

## Methods

### Objective

The objective of our case studies was to create functional AI technology-based CDS tools effective in research settings and integrate them into clinical workflow without sacrificing care quality, speed of clinical care delivery, and labor requirement.

### Setting and Participants

Cincinnati Children's Hospital Medical Center is a large tertiary care center with more than 1.2 million patient encounters

annually. It has a large epilepsy clinic (over 6,400 patients and 12,000 epilepsy visits per year) and a high volume of epilepsy surgery cases (50 per year). The division of pediatric emergency medicine oversees 5 urgent cares and 2 emergency departments (EDs) with an annual census of 170,000 visits. The ED employs 8 full-time clinical research coordinators (CRCs) who enroll patients in research studies during clinical visits.

### Case 1: Automated Epilepsy Interventions

#### Background

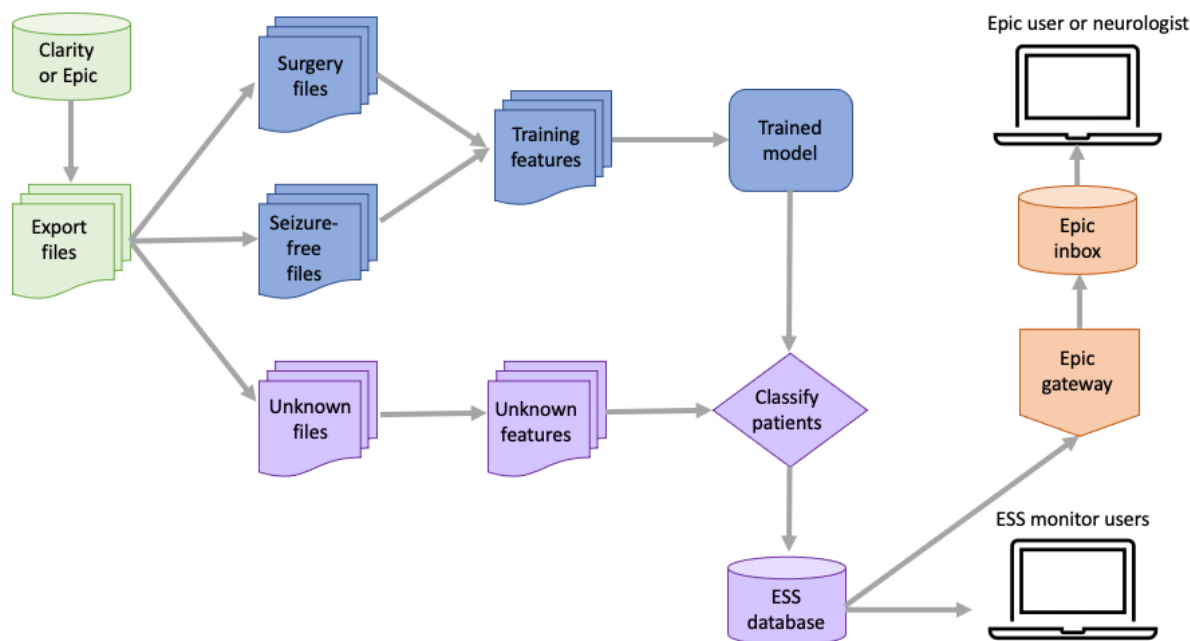
The first case study aimed to facilitate early surgical intervention in patients with intractable epilepsy, as it has been shown to improve cognitive outcomes, mental health, and quality of life [36], as well as increase quality-adjusted life years [37] in a relatively safe procedure for the patient [38]. National guidelines state that patients who continue to have debilitating seizures after 2 or more adequate trials of antiepileptic medications should be considered for a presurgical evaluation referral [39]. From the time of first seizure, on average, patients receive surgery after having epilepsy for 7 years in pediatrics and 20 years in adults [40,41]. Only 0.5%-1.5% of patients received surgery within 2 years of fulfilling clinical criteria for surgical candidacy [42]. Indeed, improving the use of surgery has proven to be difficult [42] because this highly specialized but critical clinical knowledge is not ubiquitously available in clinical care.

#### Approach

A corpus of notes from patients with a diagnosis of epilepsy who were seizure free or had a history of resective epilepsy surgery was used to devise NLP features. The NLP generated surgical candidacy scores for each patient, with higher scores indicating a higher likelihood of surgical candidacy and lower scores indicating a higher likelihood of seizure freedom. Next, naïve Bayes, support vector machine, and random forest models were developed using retrospective data as described in previous work [43]. Figure 1 describes the system pipeline from input data to the output recommendation.

To ensure the recommendations from the NLP system would be accepted into practice, we validated the algorithm's classifications by comparing them head-to-head against manual labels from epileptologists [2]. Prior to implementation into clinical care, we prospectively evaluated the system for 1 year to test the accuracy in a clinical setting [44].

**Figure 1.** Epilepsy surgical pipeline architecture. From left to right: a series of Oracle PL/SQL queries extract epilepsy patient data and export them in CSV format to bare meta installation servers. The data are divided into the following 3 groups: patients with surgery, seizure-free patients, and patients with unknown outcome. The feature extraction module (ie, ‘training features’) analyzes the free-text notes and exports machine-readable feature vectors in SVM light format. Surgery and seizure-free patient features are sent to the classifier training module to train the support vector machine model. Unknown patient features are fed into the final trained classifier, which outputs a surgery candidacy score for each patient. All patients with unknown outcome and their scores are then loaded into the Epilepsy Surgery Software (ESS) database. The highest scoring patients are sent to an Epic web service that generates the in-basket message alerts. All patients and their notes can be viewed and searched in the ESS web application. This entire process is run on a weekly basis, to continually incorporate new electronic health record data into the algorithm training.



## Case 2: Automated Clinical Trial Eligibility Screener (ACTES)

### Background

The second case study aimed to identify participants who may meet eligibility criteria for clinical trial recruitment in the ED. In current practice, CRCs and physicians at the site of the hospitals do trial eligibility screening manually [45]. For patients presenting during clinical visits, screening would ideally take place early enough in the visit such that eligible candidates could be approached for enrollment without prolonging their length of stay. However, given the large volume of data documented in EHRs, it is labor-intensive for the staff to screen relevant information, particularly within the time frame of a single visit. As such, automatically screening and identifying eligible patients for a trial based on EHR information promises great benefits for clinical research.

### Approach

To facilitate participant identification, we developed a machine learning NLP-based system—ACTES [23,46]—which analyzed structured data and unstructured narrative notes automatically to determine patients’ suitability for clinical trial enrollment. For development, we evaluated historical trial-patient enrollment decisions in a pediatric ED and extracted EHR data including clinical notes that were commonly reviewed by CRCs. We then customized the machine learning and NLP algorithms based on the trial-patient data. The ACTES was integrated into the institutional workflow to support real-time patient screening in

our recent work [44]; details of system development have been previously reported [46].

### Implementation Strategy

We hypothesized that successful implementation of the AI solutions relied on 5 key steps, as follows:

1. Integration of industry standard software pertinent to the implementation site. Specifically, the systems needed to be adapted to use industry standard software libraries.
2. Automation of the process to access the EHR data. The systems need to be linked to the EHR to extract the input data without manual intervention.
3. Encouragement of user feedback to inform the final design of the AI solution.
4. Integration of the AI solutions into typical clinical workflow.
5. Performance evaluations and regular maintenance to continue to evaluate the utility of the AI solution.

After building the AI technology, we implemented the AI solutions using these 5 strategies to facilitate successful deployment of the tools.

## Results

After creation and validation of the algorithms in a research setting, we implemented these 2 AI solutions as NLP pipelines. Both pipelines follow a step-by-step process that extracts data from the EHR, processes it, and provides a recommendation in the form of automated alerts that could be sent from the research

systems to the EHR (Epic Systems) in real time. To do this, the research systems had to be modified to integrate into clinical workflow, as described in this section.

### Industry Standard Software

After reviewing over 20 different libraries for managing NLP pipelines, it was decided that the Java NLP library LingPipe [47] would be used for feature extraction and preprocessing, and the LIBSVM Python implementation from scikit-learn [48] would be used for the classifier [49]. The NLP component in ACTES was built upon the clinical Text Analysis and Knowledge Extraction System [50], and the machine learning component was coded in Java (Oracle Corporation).

### Automation of EHR Data Access

For the epilepsy intervention AI, Oracle PL/SQL queries from the EHR relational database were used to extract patient data. For ACTES, RESTful and SOAP web services were developed to extract EHR data, such as demographics, medication orders, and clinical notes in real time, which were stored in an Oracle SQL database. An interactive web-based dashboard was developed to visualize the recommendations and receive feedback from CRCs.

### User Feedback Informed the Final Design

AI solutions were designed and integrated with feedback from end users. The epilepsy and ACTES corpora were created by manual annotation of patient notes by providers. Throughout the algorithm design and implementation process, providers were included in the build and ultimate integration. First, the biomedical informatics team shadowed providers for workflow observation. Second, the biomedical informatics team attended clinical meetings that included faculty, staff, and clinical research coordinators for a minimum of 10 hours to get feedback and ensure the design was appropriate. Third, mock-up designs were shared at a minimum of 3 meetings to discuss the process of using and interacting with the AI solution in the form of a CDS tool. In cases where the CDS tool could provide an alert, the providers were consulted on their preferred alert method (eg, email or text message alerts). In both AI technologies, the providers were able to directly interact with the machine learning recommendations as follows:

- For epilepsy surgical intervention, these results are displayed in clinical care to suggest surgical consults, and the subsequent actions resulting from the recommendations are fed back into the application to improve performance.
- For ACTES, the clinical research coordinators' entry of eligibility is used to help train and improve the classifier. Additionally, ACTES was assessed and improved for usability and satisfaction by providers and was found to be easy to use and learn.

### Integration Into Clinical Workflow

Both AI technologies were integrated into clinical workflow to support clinical practice. For patients with intractable epilepsy and an upcoming visit, surgical eligibility is evaluated in advance. For patients who are classified as potential surgical candidates, EHR in-basket messages are sent to the provider they are scheduled to see via web services.

We integrated the ACTES into the CRCs' workflow to support real-time patient screening [51]. The system ran continuously on a secured, The Health Insurance Portability and Accountability Act (HIPAA)-compliant server to extract structured and unstructured EHR data for current ED patients. For each clinical trial, the ranked list of patients recommended by the system, along with their demographics and clinical information, were displayed on the dashboard available to the CRCs. The information was refreshed at 10-minute increments to accommodate real-time updates. Given the recommended patients as potential participants for a clinical trial, the CRCs performed additional EHR screening to confirm the candidates' eligibility. When an eligible candidate was identified, the CRCs approached him or her for enrollment as per standard clinical workflow.

### Performance Evaluation

The epilepsy AI technology went live on April 12, 2016, as part of the EHR release cycle and runs weekly. On Sundays, the system trains on notes from patients who have been seizure free for 1 year or previously underwent resective epilepsy surgery. This trained classifier evaluates all other 'unknown' patients with epilepsy who have had at least one seizure within the last year but have not had a presurgical evaluation. Thus, the tables of training and test patients are updated weekly. The system performs as well as board-certified neurologists in identifying surgical candidates (with a sensitivity of 71% and positive predictive value of 77%) and improves with additional training, identifying surgical candidates faster than neurologists [2]. As part of the ongoing algorithmic development, the number of patients with a history of surgery included in the training set increased from 102 patients on April 10, 2016, to 195 patients on October 6, 2019.

The ACTES patient identification system went live on October 1, 2017. ACTES was prospectively evaluated using a time-and-motion study, quantitative assessments of enrollment, and postevaluation usability surveys collected from the CRCs [52]. During the time-and-motion study, an observer monitored the activities a CRC was engaged in at 30-second increments for 2 hours. The time spent per activity was compared to that prior to the use of ACTES. This study was repeated monthly for 4 months, and it was distributed among CRCs and shifts. After the implementation of ACTES, the CRCs spent 12.9% ( $P<.001$ ) less time on electronic screening [52]. The quantitative assessments of enrollment evaluated the number of patients screened, the number of patients approached, and the number of patients enrolled. The use of ACTES significantly improved the number of screened patients for the majority of trials and improved the number of approached patients and enrolled patients, with statistical significance in 2 of 7 trials [52]. Finally, results from the System Usability Survey and additional open-ended questions were analyzed on a monthly basis to improve ACTES [52].

### Maintenance

The epilepsy system was operational more than 90% of the time through the first 150 weeks. Throughout this time, issues were addressed by the biomedical informatics research and production IT staff. There were 10 changes made to the NLP system and

6 errors executing the pipeline of scripts. Issues extracting patient notes from the EHR were the largest reason for delays in running the NLP system, which occurred 12 out of 150 (8%) weeks.

Miscellaneous adjustments were made to the ACTES tool during the pilot phase (2017-2018) to accommodate CRC needs. ACTES was also updated 3 times because of significant updates on the institutional EHR system and its web services for real-time data extraction. Updates on the institutional EHR system and the research IT environment caused multiple system breakdowns during the evaluation period that interrupted less than 2 out of 52 (4%) weeks of operation.

## Discussion

### Principal Results

This work highlighted the major modifications for the integration and deployment of CDS tools from the research setting to clinical practice. We successfully added AI-based technology to the following 2 distinct clinical workflows at our institution: an automated epilepsy surgical intervention tool and an automated clinical trial eligibility screener (ACTES) system. Throughout the process, we determined that successful integration of these tools into clinical care requires adaptation to industry standards, automation of data access, logical integration into clinical workflow, and continual user feedback.

This work has several important strengths. We implemented novel, automated machine learning tools to provide decision support in a tangible fashion at our institution. These tools were well received and streamlined clinical care in the identification of qualified patients for surgery or clinical trials. Our experience with the deployment of these tools agreed with the suggestions made by Kawamoto et al [53] for successful implementation of CDS tools. Our CDS tools were implemented in real time to provide support at a natural point in the clinical workflow, so as not to disrupt or extend the timeline of care. As with their findings, our CDS tools use automatically available EHR data, where possible, to ensure clinical scalability and effective usability. In our case, we added an extra layer of testing whereby we implemented our CDS tools in a localized clinical setting in parallel to clinical care to test accuracy prior to full deployment, which allowed for continued fine-tuning of the CDS tool before it became part of clinical workflow.

### Acknowledgments

The authors would like to thank the neurology providers, emergency medicine providers, and biomedical informatics staff at Cincinnati Children's Hospital Medical Center for providing countless feedback and support.

This work was supported by internal funds from Cincinnati Children's Hospital Medical Center, the National Library of Medicine (5R01LM011124-03), and the Agency for Healthcare Research and Quality (1R21HS024977-01). BW's work was supported by the Ruth L. Kirschstein Predoctoral Individual National Research Service Award (5F31NS115447-02).

### Conflicts of Interest

JP and TG report a patent pending on the identification of surgical candidates using natural language processing, licensed to Cincinnati Children's Hospital Medical Center and a patent pending on processing clinical text with domain-specific spreading activation methods, licensed to Cincinnati Children's Hospital Medical Center.

### Evaluation of Bias

We evaluated both tools for potential bias to ensure that the CDS recommendations were not influenced by racial disparities. The AI technology behind epilepsy surgical candidacy recommendation was evaluated for bias in terms of patient demographics, socioeconomic characteristics, and language [54]. Patient race, gender, and primary language did not bias the AI's surgical candidacy scores ( $P > .35$  for all).

### Considerations and Limitations

Several concerns should be considered in the implementation of a research tool into real-time clinical settings. As with most record keeping systems, the EHR systems require regular upgrades and bug fixes. This necessitates ongoing IT support to keep the pipeline operational. EHR algorithm extractions and pipeline characteristics should be placed into the EHR upgrade queue to ensure their evaluation with each upgrade cycle. To account for this, resources for both operational and research IT should be set aside to ensure a consistent system when integrated with clinical practice.

The successful deployment and continued use of these systems also required close collaboration with the stakeholders embedded in the respective clinical system. This collaboration was crucial in allowing seamless integration of the research output into daily clinical practice. Without input from the effective end users, it would be difficult to fully understand the current process, needs, as well as limitations related to workflow and data to allow for optimization of the prediction.

### Conclusions

The formulation, development, and real-time implementation of two AI solutions in a clinical setting required the development of a CDS tool and pipeline using public, industry-standard programs and existing EHR web interfaces prior to integration. In our work, we found that a CDS tool's success was largely dependent upon the collaboration between machine learning experts, research collaborators, and operational IT professionals. Furthermore, longitudinal support from clinical providers and institutional leadership is necessary for continued maintenance of the deployed CDS tool with careful consideration for its long-term use.

## References

1. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019 Apr 04;380(14):1347-1358. [doi: [10.1056/nejmra1814259](https://doi.org/10.1056/nejmra1814259)]
2. Cohen KB, Glass B, Greiner HM, Holland-Bouley K, Standridge S, Arya R, et al. Methodological issues in predicting pediatric epilepsy surgery candidates through natural language processing and machine learning. *Biomed Inform Insights* 2016 May 22;8:BII.S38308. [doi: [10.4137/bii.s38308](https://doi.org/10.4137/bii.s38308)]
3. Matykiewicz P, Cohen K, Holland KD, Glauser TA, Standridge SM, Verspoor KM, et al. Earlier identification of epilepsy surgery candidates using natural language processing. 2013 Presented at: Proceedings of the 2013 Workshop on Biomedical Natural Language Processing; August 8; Sofia, Bulgaria p. 1-9.
4. Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schragger JD. Prediction of emergency department hospital admission based on natural language processing and neural networks. *Methods Inf Med* 2017 Oct 26;56(5):377-389. [doi: [10.3414/ME17-01-0024](https://doi.org/10.3414/ME17-01-0024)] [Medline: [28816338](https://pubmed.ncbi.nlm.nih.gov/28816338/)]
5. Zeng Z, Shi H, Wu Y, Hong Z. Survey of natural language processing techniques in bioinformatics. *Comput Math Methods Med* 2015;2015:674296-674210 [FREE Full text] [doi: [10.1155/2015/674296](https://doi.org/10.1155/2015/674296)] [Medline: [26525745](https://pubmed.ncbi.nlm.nih.gov/26525745/)]
6. Milea D, Najjar RP, Jiang Z, Ting D, Vasseneix C, Xu X, et al. Artificial intelligence to detect papilledema from ocular fundus photographs. *N Engl J Med* 2020 Apr 30;382(18):1687-1695. [doi: [10.1056/nejmoa1917130](https://doi.org/10.1056/nejmoa1917130)]
7. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
8. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019 Aug 31;572(7767):116-119 [FREE Full text] [doi: [10.1038/s41586-019-1390-1](https://doi.org/10.1038/s41586-019-1390-1)] [Medline: [31367026](https://pubmed.ncbi.nlm.nih.gov/31367026/)]
9. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017 Jan 25;542(7639):115-118. [doi: [10.1038/nature21056](https://doi.org/10.1038/nature21056)]
10. Hicks J. The potential of claims data to support the measurement of health care quality. RAND Corporation. 2003. URL: [https://www.rand.org/pubs/rgs\\_dissertations/RGSD171.html](https://www.rand.org/pubs/rgs_dissertations/RGSD171.html) [accessed 2022-11-17]
11. Hirschberg J, Manning CD. Advances in natural language processing. *Science* 2015 Jul 17;349(6245):261-266. [doi: [10.1126/science.aaa8685](https://doi.org/10.1126/science.aaa8685)] [Medline: [26185244](https://pubmed.ncbi.nlm.nih.gov/26185244/)]
12. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc* 2005 Jul 01;12(4):448-457. [doi: [10.1197/jamia.m1794](https://doi.org/10.1197/jamia.m1794)]
13. Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* 2011 Aug 24;306(8):848-855. [doi: [10.1001/jama.2011.1204](https://doi.org/10.1001/jama.2011.1204)] [Medline: [21862746](https://pubmed.ncbi.nlm.nih.gov/21862746/)]
14. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J Am Med Inform Assoc* 2003 Mar 01;10(2):115-128 [FREE Full text] [doi: [10.1197/jamia.m1074](https://doi.org/10.1197/jamia.m1074)] [Medline: [12595401](https://pubmed.ncbi.nlm.nih.gov/12595401/)]
15. Petratos GN, Kim Y, Evans RS, Williams SD, Gardner RM. Comparing the effectiveness of computerized adverse drug event monitoring systems to enhance clinical decision support for hospitalized patients. *Appl Clin Inform* 2017 Dec 16;01(03):293-303. [doi: [10.4338/aci-2009-11-ra-0009](https://doi.org/10.4338/aci-2009-11-ra-0009)]
16. Tinoco A, Evans RS, Staes CJ, Lloyd JF, Rothschild JM, Haug PJ. Comparison of computerized surveillance and manual chart review for adverse events. *J Am Med Inform Assoc* 2011 Jul 01;18(4):491-497 [FREE Full text] [doi: [10.1136/amiajnl-2011-000187](https://doi.org/10.1136/amiajnl-2011-000187)] [Medline: [21672911](https://pubmed.ncbi.nlm.nih.gov/21672911/)]
17. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: a review of chief complaint based classifiers in North America. *J Biomed Inform* 2013 Aug;46(4):734-743 [FREE Full text] [doi: [10.1016/j.jbi.2013.04.003](https://doi.org/10.1016/j.jbi.2013.04.003)] [Medline: [23602781](https://pubmed.ncbi.nlm.nih.gov/23602781/)]
18. Ye Y, Tsui F, Wagner M, Espino JU, Li Q. Influenza detection from emergency department reports using natural language processing and Bayesian network classifiers. *J Am Med Inform Assoc* 2014 Sep 01;21(5):815-823 [FREE Full text] [doi: [10.1136/amiajnl-2013-001934](https://doi.org/10.1136/amiajnl-2013-001934)] [Medline: [24406261](https://pubmed.ncbi.nlm.nih.gov/24406261/)]
19. Savova GK, Olson JE, Murphy SP, Cafourek VL, Couch FJ, Goetz MP, et al. Automated discovery of drug treatment patterns for endocrine therapy of breast cancer within an electronic medical record. *J Am Med Inform Assoc* 2012 Jun 01;19(e1):e83-e89 [FREE Full text] [doi: [10.1136/amiajnl-2011-000295](https://doi.org/10.1136/amiajnl-2011-000295)] [Medline: [22140207](https://pubmed.ncbi.nlm.nih.gov/22140207/)]
20. Dublin S, Baldwin E, Walker RL, Christensen LM, Haug PJ, Jackson ML, et al. Natural Language Processing to identify pneumonia from radiology reports. *Pharmacoepidemiol Drug Saf* 2013 Aug 01;22(8):834-841 [FREE Full text] [doi: [10.1002/pds.3418](https://doi.org/10.1002/pds.3418)] [Medline: [23554109](https://pubmed.ncbi.nlm.nih.gov/23554109/)]
21. Elkin PL, Froehling D, Wahner-Roedler D, Trusko B, Welsh G, Ma H, et al. NLP-based identification of pneumonia cases from free-text radiological reports. *AMIA Annu Symp Proc* 2008 Nov 06:172-176 [FREE Full text] [Medline: [18998791](https://pubmed.ncbi.nlm.nih.gov/18998791/)]
22. Friedman C, Alderson PO, Austin JHM, Cimino JJ, Johnson SB. A general natural-language text processor for clinical radiology. *J Am Med Inform Assoc* 1994 Mar 01;1(2):161-174 [FREE Full text] [doi: [10.1136/jamia.1994.95236146](https://doi.org/10.1136/jamia.1994.95236146)] [Medline: [7719797](https://pubmed.ncbi.nlm.nih.gov/7719797/)]

23. Deleger L, Brodzinski H, Zhai H, Li Q, Lingren T, Kirkendall ES, et al. Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. *J Am Med Inform Assoc* 2013 Dec 01;20(e2):e212-e220 [FREE Full text] [doi: [10.1136/amiajnl-2013-001962](https://doi.org/10.1136/amiajnl-2013-001962)] [Medline: [24130231](https://pubmed.ncbi.nlm.nih.gov/24130231/)]
24. Branch-Elliman W, Strymish J, Kudesia V, Rosen AK, Gupta K. Natural language processing for real-time catheter-associated urinary tract infection surveillance: results of a pilot implementation trial. *Infect Control Hosp Epidemiol* 2015 Sep 26;36(9):1004-1010. [doi: [10.1017/ice.2015.122](https://doi.org/10.1017/ice.2015.122)] [Medline: [26022228](https://pubmed.ncbi.nlm.nih.gov/26022228/)]
25. Tso GJ, Tu SW, Oshiro C, Martins S, Ashcraft M, Yuen KW, et al. Automating guidelines for clinical decision support: knowledge engineering and implementation. *AMIA Annu Symp Proc* 2016;2016:1189-1198 [FREE Full text] [Medline: [28269916](https://pubmed.ncbi.nlm.nih.gov/28269916/)]
26. Klein ME, Parvez MM, Shin J. Clinical implementation of pharmacogenomics for personalized precision medicine: barriers and solutions. *J Pharm Sci* 2017 Sep;106(9):2368-2379. [doi: [10.1016/j.xphs.2017.04.051](https://doi.org/10.1016/j.xphs.2017.04.051)] [Medline: [28619604](https://pubmed.ncbi.nlm.nih.gov/28619604/)]
27. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: a systematic review and gaps analysis. *Int J Med Inform* 2017 Dec;98:56-64. [doi: [10.1016/j.ijmedinf.2016.12.001](https://doi.org/10.1016/j.ijmedinf.2016.12.001)] [Medline: [28034413](https://pubmed.ncbi.nlm.nih.gov/28034413/)]
28. Castillo RS, Kelemen A. Considerations for a successful clinical decision support system. *Comput Inform Nurs* 2013 Jul;31(7):319-26; quiz 327-8. [doi: [10.1097/NXN.0b013e3182997a9c](https://doi.org/10.1097/NXN.0b013e3182997a9c)] [Medline: [23774450](https://pubmed.ncbi.nlm.nih.gov/23774450/)]
29. Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005 Mar 9;293(10):1223-1238. [doi: [10.1001/jama.293.10.1223](https://doi.org/10.1001/jama.293.10.1223)] [Medline: [15755945](https://pubmed.ncbi.nlm.nih.gov/15755945/)]
30. Wright A, Ash JS, Aaron S, Ai A, Hickman TT, Wiesen JF, et al. Best practices for preventing malfunctions in rule-based clinical decision support alerts and reminders: results of a Delphi study. *Int J Med Inform* 2018 Oct;118:78-85 [FREE Full text] [doi: [10.1016/j.ijmedinf.2018.08.001](https://doi.org/10.1016/j.ijmedinf.2018.08.001)] [Medline: [30153926](https://pubmed.ncbi.nlm.nih.gov/30153926/)]
31. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018 Aug 28;1(1):39 [FREE Full text] [doi: [10.1038/s41746-018-0040-6](https://doi.org/10.1038/s41746-018-0040-6)] [Medline: [31304320](https://pubmed.ncbi.nlm.nih.gov/31304320/)]
32. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat Med* 2020 Jan;26(1):52-58 [FREE Full text] [doi: [10.1038/s41591-019-0715-9](https://doi.org/10.1038/s41591-019-0715-9)] [Medline: [31907460](https://pubmed.ncbi.nlm.nih.gov/31907460/)]
33. Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med* 2018 Sep 13;24(9):1337-1341. [doi: [10.1038/s41591-018-0147-y](https://doi.org/10.1038/s41591-018-0147-y)] [Medline: [30104767](https://pubmed.ncbi.nlm.nih.gov/30104767/)]
34. Wang P, Liu X, Berzin T, Glissen Brown J, Liu P, Zhou C, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020 Apr;5(4):343-351 [FREE Full text] [doi: [10.1016/s2468-1253\(19\)30411-x](https://doi.org/10.1016/s2468-1253(19)30411-x)]
35. Trinkley KE, Kahn MG, Bennett TD, Glasgow RE, Haugen H, Kao DP, et al. Integrating the practical robust implementation and sustainability model with best practices in clinical decision support design: implementation science approach. *J Med Internet Res* 2020 Oct 29;22(10):e19676 [FREE Full text] [doi: [10.2196/19676](https://doi.org/10.2196/19676)] [Medline: [33118943](https://pubmed.ncbi.nlm.nih.gov/33118943/)]
36. Engel J, McDermott MP, Wiebe S, Langfitt JT, Stern JM, Dewar S, Early Randomized Surgical Epilepsy Trial (ERSET) Study Group. Early surgical therapy for drug-resistant temporal lobe epilepsy: a randomized trial. *JAMA* 2012 Mar 07;307(9):922-930 [FREE Full text] [doi: [10.1001/jama.2012.220](https://doi.org/10.1001/jama.2012.220)] [Medline: [22396514](https://pubmed.ncbi.nlm.nih.gov/22396514/)]
37. Choi H, Sell RL, Lenert L, Muennig P, Goodman RR, Gilliam FG, et al. Epilepsy surgery for pharmacoresistant temporal lobe epilepsy: a decision analysis. *JAMA* 2008 Dec 03;300(21):2497-2505. [doi: [10.1001/jama.2008.771](https://doi.org/10.1001/jama.2008.771)] [Medline: [19050193](https://pubmed.ncbi.nlm.nih.gov/19050193/)]
38. Engel J, Wiebe S, French J, Sperling M, Williamson P, Spencer D, et al. Practice parameter: temporal lobe and localized neocortical resections for epilepsy. *Epilepsia* 2003 Jun;44(6):741-751 [FREE Full text] [doi: [10.1046/j.1528-1157.2003.48202.x](https://doi.org/10.1046/j.1528-1157.2003.48202.x)] [Medline: [12790886](https://pubmed.ncbi.nlm.nih.gov/12790886/)]
39. Cross JH, Jayakar P, Nordli D, Delalande O, Duchowny M, Wieser HG, International League against Epilepsy, Subcommission for Paediatric Epilepsy Surgery, Commissions of NeurosurgeryPaediatrics. Proposed criteria for referral and evaluation of children for epilepsy surgery: recommendations of the Subcommission for Pediatric Epilepsy Surgery. *Epilepsia* 2006 Jun;47(6):952-959 [FREE Full text] [doi: [10.1111/j.1528-1167.2006.00569.x](https://doi.org/10.1111/j.1528-1167.2006.00569.x)] [Medline: [16822241](https://pubmed.ncbi.nlm.nih.gov/16822241/)]
40. Choi H, Carlino R, Heiman G, Hauser WA, Gilliam FG. Evaluation of duration of epilepsy prior to temporal lobe epilepsy surgery during the past two decades. *Epilepsy Res* 2009 Oct;86(2-3):224-227 [FREE Full text] [doi: [10.1016/j.epilepsyres.2009.05.014](https://doi.org/10.1016/j.epilepsyres.2009.05.014)] [Medline: [19581072](https://pubmed.ncbi.nlm.nih.gov/19581072/)]
41. Kwan P, Schachter SC, Brodie MJ. Drug-resistant epilepsy. *N Engl J Med* 2011 Sep 08;365(10):919-926. [doi: [10.1056/nejmra1004418](https://doi.org/10.1056/nejmra1004418)]
42. Englot DJ, Ouyang D, Garcia PA, Barbaro NM, Chang EF. Epilepsy surgery trends in the United States, 1990-2008. *Neurology* 2012 Mar 21;78(16):1200-1206. [doi: [10.1212/wnl.0b013e318250d7ea](https://doi.org/10.1212/wnl.0b013e318250d7ea)]

43. Tsochantaridis I, Hofmann T, Thorsten J, Altun Y. Support vector machine learning for interdependent and structured output spaces. In: Proceedings of the twenty-first international conference on Machine learning. 2004 Presented at: ICML '04; July 4-8; Banff, Alberta, Canada. [doi: [10.1145/1015330.1015341](https://doi.org/10.1145/1015330.1015341)]
44. Wissel BD, Greiner HM, Glauser TA, Holland-Bouley KD, Mangano FT, Santel D, et al. Prospective validation of a machine learning model that uses provider notes to identify candidates for resective epilepsy surgery. *Epilepsia* 2020 Jan 29;61(1):39-48 [FREE Full text] [doi: [10.1111/epi.16398](https://doi.org/10.1111/epi.16398)] [Medline: [31784992](https://pubmed.ncbi.nlm.nih.gov/31784992/)]
45. Embi PJ, Payne PRO. Clinical research informatics: challenges, opportunities and definition for an emerging domain. *JAMIA* 2009 May 01;16(3):316-327. [doi: [10.1197/jamia.m3005](https://doi.org/10.1197/jamia.m3005)]
46. Ni Y, Kennebeck S, Dexheimer JW, McAnaney CM, Tang H, Lingren T, et al. Automated clinical trial eligibility prescreening: increasing the efficiency of patient identification for clinical trials in the emergency department. *J Am Med Inform Assoc* 2015 Jan;22(1):166-178 [FREE Full text] [doi: [10.1136/amiajnl-2014-002887](https://doi.org/10.1136/amiajnl-2014-002887)] [Medline: [25030032](https://pubmed.ncbi.nlm.nih.gov/25030032/)]
47. Baldwin B, Dayanidhi K. Natural language processing with Java and LingPipe Cookbook. Birmingham, UK: Packt Publishing Ltd; 2014.
48. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;12:2825-2830. [doi: [10.5555/1953048.2078195](https://doi.org/10.5555/1953048.2078195)]
49. Joachims T. Text categorization with support vector machines: learning with many relevant features. : Springer Berlin Heidelberg; 1998 Presented at: Machine Learning: ECML-98; 1998; Berlin, Heidelberg p. 137-142. [doi: [10.1007/BFb0026683](https://doi.org/10.1007/BFb0026683)]
50. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc* 2010;17(5):507-513 [FREE Full text] [doi: [10.1136/jamia.2009.001560](https://doi.org/10.1136/jamia.2009.001560)] [Medline: [20819853](https://pubmed.ncbi.nlm.nih.gov/20819853/)]
51. Dexheimer JW, Tang H, Kachelmeyer A, Houchell M, Kennebeck S, Solti I, et al. A time-and-motion study of clinical trial eligibility screening in a pediatric emergency department. *Pediatr Emerg Care* 2019 Dec;35(12):868-873 [FREE Full text] [doi: [10.1097/PEC.0000000000001592](https://doi.org/10.1097/PEC.0000000000001592)] [Medline: [30281551](https://pubmed.ncbi.nlm.nih.gov/30281551/)]
52. Ni Y, Bermudez M, Kennebeck S, Liddy-Hicks S, Dexheimer J. A real-time automated patient screening system for clinical trials eligibility in an emergency department: design and evaluation. *JMIR Med Inform* 2019 Jul 24;7(3):e14185 [FREE Full text] [doi: [10.2196/14185](https://doi.org/10.2196/14185)] [Medline: [31342909](https://pubmed.ncbi.nlm.nih.gov/31342909/)]
53. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005 Apr 2;330(7494):765 [FREE Full text] [doi: [10.1136/bmj.38398.500764.8F](https://doi.org/10.1136/bmj.38398.500764.8F)] [Medline: [15767266](https://pubmed.ncbi.nlm.nih.gov/15767266/)]
54. Wissel BD, Greiner HM, Glauser TA, Mangano FT, Santel D, Pestian JP, et al. Investigation of bias in an epilepsy machine learning algorithm trained on physician notes. *Epilepsia* 2019 Sep 23;60(9):e93-e98 [FREE Full text] [doi: [10.1111/epi.16320](https://doi.org/10.1111/epi.16320)] [Medline: [31441044](https://pubmed.ncbi.nlm.nih.gov/31441044/)]

## Abbreviations

**ACTES:** automated clinical trial eligibility screener  
**AI:** artificial intelligence  
**CDS:** clinical decision support  
**CRC:** clinical research coordinator  
**ED:** emergency department  
**EHR:** electronic health record  
**IT:** information technology  
**NLP:** natural language processing

*Edited by T Hao; submitted 08.03.22; peer-reviewed by A Keniston, H Monday, G Nneji; comments to author 11.07.22; revised version received 01.09.22; accepted 19.09.22; published 16.12.22*

### *Please cite as:*

Kanbar LJ, Wissel B, Ni Y, Pajor N, Glauser T, Pestian J, Dexheimer JW  
*Implementation of Machine Learning Pipelines for Clinical Practice: Development and Validation Study*  
*JMIR Med Inform* 2022;10(12):e37833  
URL: <https://medinform.jmir.org/2022/12/e37833>  
doi: [10.2196/37833](https://doi.org/10.2196/37833)  
PMID:



©Lara J Kanbar, Benjamin Wissel, Yizhao Ni, Nathan Pajor, Tracy Glauser, John Pestian, Judith W Dexheimer. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 16.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.