

Findings from the 2019 International Medical Informatics Association Yearbook Section on Health Information Management

Meryl Bloomrosen¹, Eta S. Berner², Section Editors for the IMIA Yearbook Section on Health Information Management

¹ Premier healthcare alliance, Washington, DC, USA

² Graduate Programs in Health Informatics, Department of Health Services Administration, University of Alabama at Birmingham, Birmingham, AL, USA

Summary

Objectives: To summarize the recent literature and research and present a selection of the best papers published in 2018 in the field of Health Information Management (HIM) and Health Informatics.

Methods: A systematic review of the literature was performed, with the help of a medical librarian, by the two editors of the HIM section of the International Medical Informatics Association (IMIA) Yearbook. In order to include papers that would address the special theme of the 2019 Yearbook on artificial intelligence (AI) as well as HIM, we searched bibliographic databases for HIM-related papers with an AI focus using both Medical Subject Headings (MeSH) descriptors and keywords in titles and abstracts. A shortlist of 15 candidate best papers was first selected by section editors before being peer-reviewed by independent external reviewers.

Results: While there were a significant number of manuscripts that addressed issues relevant to HIM, there were virtually none with MeSH headings indicating an HIM focus. Manuscripts that were considered related to the HIM field in terms of the practice of health information management as well as the profession

included those that examined using machine learning and other AI approaches to identify protected health information in clinical text to aid with de-identification, automated coding approaches to translate free-text into standardized codes, and natural language processing approaches to extract clinical data to assist with populating cancer and other registries.

Conclusions: The papers discussed in the HIM section reflect the special theme of the use of AI in healthcare on issues particularly relevant to the field of HIM. This synopsis discusses these papers and recommends that HIM practitioners be more involved in research and that researchers in AI and related areas recognize the applicability and relevance of their work to the field of HIM.

Keywords

Health information management; artificial intelligence; automated coding; text mining, natural language processing; electronic health records

Yerb Med Inform 2019;65-70

<http://dx.doi.org/10.1055/s-0039-1677941>

can be used to identify not only protected health information (PHI) for de-identification purposes, but also other specific elements in clinical text that can assist in other clinical and research tasks. While taking the responsibility for curating information, or maintaining registries, for example, are key roles for HIM professionals, there are other roles within and beyond HIM that can also benefit from the use of AI-related techniques.

This synopsis looks at the literature published in 2018 on AI approaches used in contexts and for purposes of specific interest and relevance to the field of HIM. With the evolution of HIM, increased digitization of health data, and more evidence of the potential benefits of AI applications, there will likely be ongoing adoption and implementation of AI to other administrative, clinical, and operational processes related to HIM, requiring further study and evaluation. Areas for potential future focus may include revenue cycle management, clinical trials recruitment, predictive analytics, documentation review, claims adjudication and processing, and population health management.

1 Introduction

As electronic health records (EHRs) have become widespread in healthcare, the field of Health Information Management (HIM) has evolved to include responsibility for the management of health information in electronic formats. With the vast amount of information now in digital form, there is increasing interest in using advanced algorithms, machine learning, and other forms of artificial intelligence (AI) approaches to mine the records for a variety of research and operational purposes. HIM professionals have

begun, and they will need to continue, to use these approaches to make their work as the stewards of health information more reliable and efficient. For example, automated coding software has been used in US hospitals for a number of years to assist HIM professionals in assigning diagnostic and procedure codes [1]. Protecting the privacy of health information with advanced de-identification techniques is becoming necessary, as there is an increasing amount of secondary use of the data in EHRs. Finally, many of these AI techniques and specifically natural language processing (NLP) and machine learning techniques and methods

2 Methods

In January 2019, with the assistance of a medical librarian, the editors of the HIM section of the International Medical Informatics Association (IMIA) Yearbook conducted a search of both PubMed and Embase using both Medical Subject Headings (MeSH) descriptors and keywords in the titles and

abstracts. Our intention was to focus on Health Information Management and from the articles identified, to select for review those that had a focus on AI or related concepts. However, we added a search that also related to automated coding to pick up other relevant articles. The publication year was 2018 and did not include those articles that were e-published ahead of print. The first query for PubMed was: “Health Information Management”[Mesh] OR “Health Information Management”[tiab] OR “HIM J”[Journal] OR “JAHIMA”[Journal], which yielded 98 results. The second search was “((Automation [mesh] OR automation [tiab]) AND (“Clinical Coding”[Mesh] OR coding [tiab]))” which yielded 15 results.

The first query for EMBASE was ‘medical information system’/exp/mj OR “health information management”:ti,ab OR “clinical information system”:ti,ab OR “clinical pharmacy information systems”:ti,ab OR “health information exchange”:ti,ab OR “health information management”:ti,ab OR “health information manager”:ti,ab OR “health information network”:ti,ab OR “health information system”:ti,ab OR “health information systems”:ti,ab OR “IS-H med”:ti,ab OR “medical information service”:ti,ab OR ‘Health Information Management Journal’ which, yielded 152 non-duplicative articles. The second EMBASE search was ((‘automation’/exp OR automation:ti,ab OR automatization:ti,ab OR computerization:ti,ab) AND (‘coding’/exp OR coding:ti,ab OR “information codification”:ti,ab)), which found 19 unique articles. The total of 284 articles was reviewed for appropriate articles focusing on AI and related concepts. Unfortunately, very few relevant articles were found. Therefore, we conducted another search that did not include HIM keywords or MeSH headings, but which focused directly on AI and EHRs. The new search strategy for PubMed was („Artificial Intelligence“[Mesh] OR „Artificial Intelligence“[tiab] OR “Computational Intelligence” [tiab] OR “Machine Intelligence”[tiab] OR “Computer Reasoning”[tiab] OR Computer-Vision-System*[tiab] OR “machine learning”[tiab] OR “deep learning”[tiab] OR “hierarchical learning”[tiab]) AND („Electronic Health Records“[Mesh] OR electronic-health-re-

cord*[tiab] OR electronic-medical-record*[tiab] OR computerized-health-record*[tiab] OR computerized-medical-record*[tiab]). This search led to 222 unique articles.

The new search for Embase was (,artificial intelligence’/exp OR “artificial intelligence”:ti,ab OR ,machine learning’/exp OR “machine learning”:ti,ab OR “deep learning”:ti,ab OR “learning machine*”:ti,ab OR “hierarchical learning”:ti,ab OR “Computational Intelligence”:ti,ab OR “Machine Intelligence”:ti,ab OR “Computer Reasoning”:ti,ab OR “Computer Vision System*”:ti,ab) AND (,electronic health record’/exp OR “electronic health record*”:ti,ab OR “electronic medical record*”:ti,ab OR “computerized health record*”:ti,ab OR “computerized medical record*”:ti,ab), which yielded 46 unique results. Thus, the total number of articles reviewed was 552.

The 552 unique articles were rated by both section editors, who excluded articles that were opinion pieces, or editorials, or articles where the full text of the article was not readily available. Each of the two section editors independently judged the relevance to the HIM field as well as the focus on AI and related concepts, and the quality of the articles. Those that both co-editors rated as not appropriate were excluded automatically. The rest of the articles were discussed, and disagreements adjudicated to arrive at 15 articles that, based primarily on the abstracts, were judged to be of good quality and reflected diverse aspects of the special theme of this year’s Yearbook. The full texts of these 15 articles were then rated independently by both section editors, one of the Yearbook editors, and at least one external peer reviewer.

Among the 15 candidate best papers, we selected four ‘Best Papers’ based primarily on consensus of reviewers. Other factors included having a high average rating from the reviewers, diversity of research approaches or focal area, and setting diversity.

The survey paper for the HIM section focuses primarily on HIM practices and application that are likely to be affected by the growth of AI approaches [2]. These areas include privacy of health data, medical coding, data management as well as data governance, and HIM workforce development. While there is some overlap of the themes from the best papers for 2018, the survey paper is broader in terms of the literature reviewed and the time period in which the studies were conducted. Below we discuss the major themes of the 15 candidate best papers from 2018.

3 Results

3.1 Automated Identification of Data for De-identification

Natural language processing (NLP) and other types of AI approaches have been used to identify PHI in clinical text for the purpose of removing it in order to provide de-identified clinical data for research and other purposes [3]. As EHR adoption and use continue to expand to non-English speaking countries, the need to evaluate the appropriateness of NLP methods for non-English clinical documents becomes increasingly important. However, there are fewer studies on how similar methods can be

Table 1 Best paper selection of articles for the IMIA Yearbook of Medical Informatics 2019 in the section ‘Health Information Management’. The articles are listed in alphabetical order of the first author’s surname.

Section
Health Information Management
<ul style="list-style-type: none"> ▪ Atutxa A, Pérez A, Casillas A. Machine learning approaches on diagnostic term encoding with the ICD for clinical documentation. <i>IEEE J Biomed Health Inform</i> 2018;22(4):1323-9. ▪ Cui L, Xie X, Shen Z. Prediction task guided representation learning of medical codes in EHR. <i>J Biomed Inform</i> 2018;84:1-10. ▪ Li F, Liu W, Yu H. Extraction of information related to adverse drug events from electronic health record notes: design of an end-to-end model based on deep learning. <i>JMIR Med Inform</i> 2018;6(4):e12159. ▪ Qiu JX, Yoon H-J, Fearm PA, Tourassi GD. Deep learning for automated extraction of primary sites from cancer pathology reports. <i>IEEE J Biomed Health Inform</i> 2018;22(1):244-51.

used in non-English clinical texts. This type of analysis can be particularly challenging in many of the Asian countries where the format of the text itself, which does not clearly separate words, poses difficulties. Three of the papers in our set of 15 candidates used the conditional random fields (CRF) method to identify specific types of information in Chinese or Korean clinical documents. Two studies examined Chinese clinical texts [4, 5] and one examined Korean clinical data [6]. The study by Du *et al.*, [4] focused specifically on identifying PHI in Chinese clinical texts, while Zhang *et al.*, [5] tried to identify specific named clinical entities in Chinese texts. Lee *et al.*, [6] worked with Korean clinical texts. Some of the studies used only discharge summaries, while others used clinical notes. All three studies concluded that their methods had good potential to identify key pieces of information in clinical text.

3.2 Automated Coding

The use of automated coding assistance for assigning billing codes is now a routine part of HIM practice in the US. With the increased recognition of the need for greater interoperability and data sharing, standards such as RxNorm for medications and Logical Observation Identifiers Names & Codes (LOINC) for laboratory tests have been proposed [7]. Laboratory tests in particular are often described differently in different hospital systems, and there is a need to map local test names to a standard that can be used to share data for clinical and research purposes. Parr *et al.*, used machine learning for automated mapping of laboratory tests from the US Veterans' Administration to LOINC codes [8]. The authors considered their work successful and felt it could be a model to reduce the labor intensive work of manual mapping of laboratory tests to a common standard that could be used for health data exchange for clinical care or research.

In other countries, with different health-care payment systems, the use of these codes may not be tied to reimbursement per se, but since the ICD-10 (International Classification of Diseases 10th revision) is an international standard, it can be used as a standard for exchange of information on a global scale

and as structured data in analytic models for operational and research use. Unfortunately, to date, automated coding assistance for non-English clinical data is not as well developed as for English clinical text. In addition, often the data used for automated coding is based on discharge summaries or other parts of the medical record, which may not capture the full clinical text. Several of the reviewed papers address these problems. Cui *et al.*, in a paper selected as Best Paper [9], used structured data including medical codes from three years of data from five Chinese hospitals and novel methods to develop a way of aggregating data for prediction models. While the data included some standard codes, developing automated assistance in assigning codes for other non-English data sets is still a challenge. Two papers by the same research team, one of which was selected as a Best Paper [10], address this issue for clinical documentation in Spanish [10,11]. Continued research on both automated methods for assigning standard codes for non-English clinical text as well as methods to link different code sets to each other and to analytic approaches will facilitate exchange of information for both operational and research uses.

3.3 AI Methods for Cancer Registries

Many HIM professionals are responsible for reviewing clinical data, as well as laboratory, surgical, and procedure reports to identify both appropriate cases and specific data to populate cancer registries. This process can be very time-consuming, and the sheer burden of the task may lead to delays in reporting. Just as automated coding software can make the coding process more accurate and more efficient, applying NLP and other machine learning methods to identify cancer-related data in EHRs can potentially improve the efficiency of populating cancer registries, and can identify data for research purposes as well. Qiu *et al.*, in one of the Best Papers for this section, compared several methods to extract data from pathology reports to identify the primary sites for cancer [12]. This piece of information is an important part of what must be entered into cancer registries. The authors were able to identify the best performing methods and they discussed the

challenges in obtaining this type of information. Tang *et al.*, used standard machine learning methods to identify breast pathology in Chinese pathology reports [13]. Their gold standard was having physicians who were proficient in Chinese review the same reports. The standard methods performed acceptably compared to the physicians. Miao *et al.*, also examined cancer-related information in Chinese texts but they used breast ultrasound reports to identify data to meet the standards of the American College of Radiology for reporting on breast radiologic findings [14]. The authors felt that being able to apply American standards to Chinese texts will facilitate international collaborations.

The structured data in cancer registries can be a useful source of data for research, especially if it is combined with other clinical data. In addition to entering data from routine clinical reports into cancer registries, cancer registrars often are called upon to assist with the extraction and merging of data from cancer registries for other purposes. Researchers at the University of Pittsburgh developed a tool to extract and merge cancer registry data with unstructured data from other reports [15].

3.4 Extraction of Non-Cancer Related Data from Clinical Notes

Clinical research often has required time-consuming chart review to obtain the needed research data from unstructured notes. HIM professionals, with their expertise in interpreting the clinical data, are often engaged to assist with this chart review. A set of candidate best papers have used deep learning and NLP methods to assist with this extraction. The research by Li *et al.*, one of the Best Papers in this section, used deep learning methods to augment traditional NLP methods to identify adverse drug events [16]. They also compared their method to several standard approaches. Chu *et al.*, also extracted adverse event data, but they used a neural attention-based network [17]. Their model performed better than several more traditional models.

Afzal *et al.*, used NLP on clinical notes to develop an algorithm to identify critical limb ischemia [18]. Their gold standard comparison was ICD-9 codes and the researchers

found that NLP showed potential. Leroy *et al.*, focused on using NLP methods applied to clinical records to identify diagnostic criteria for Autism Spectrum Disorders (ASD) [19]. Although this research was done on a clinical dataset collected prior to the most recent version of the criteria for ASD [20] and used the earlier criteria, the approach showed potential to facilitate identification of criteria for research as well as ASD surveillance in the general population.

4 Conclusion

Almost all the papers in this review applied AI, machine learning, and NLP techniques to extract structured data from unstructured clinical narratives in both English data sources as well as sources in other languages. Tasks such as applying billing codes or populating cancer registries or assisting with clinical research are key roles for HIM professionals. Collectively, the set of papers show the potential for these techniques to improve the efficiency of what have been laborious manual processes.

In the future, the uses of AI and machine learning methods to mine structured, and increasingly, unstructured, data from EHRs are likely to expand. Such expansion, in addition to clinical and health services research that make use of data in EHRs, might also include risk scoring and other predictive modeling, population health management, analyses for revenue enhancement, and quality assurance activities. As the survey paper of the HIM section of the IMIA Yearbook, authored by Stanfill et al. [2] makes clear, when the use of these methods becomes more integrated into research and clinical activities, the need to address a variety of technical and ethical issues, including those related to data quality, as well as privacy and security, will be increasingly recognized. HIM professionals can play a key role in addressing these issues, but the issues themselves are important to many professions and multiple and diverse research domains.

Given the importance of AI methods and approaches to the field of Health Information Management, it was striking that the MeSH headings of papers that represent cutting

edge work in the use of AI concepts rarely included MeSH headings related to HIM, although these articles could be found with searches that included the AI concepts and EHRs. Similarly, the set of papers that included HIM-related MeSH headings did not include papers on AI methods. It is difficult to tell whether the lack of overlap of the AI literature and HIM is a result of how the article authors chose key words, how the MeSH coders assigned headings, or the fact that HIM professionals are not involved in this research and the researchers do not identify with HIM. Whatever the cause, the results of the 2018 literature search as well as the discussion in the survey paper highlight the need for HIM professionals to become more knowledgeable about these new approaches and to bring their expertise to the research applying these methods in practice.

Acknowledgement

We would like to acknowledge Megan Bell, who assisted us with formulating our search strategy. We also appreciate the guidance and support of the entire Yearbook editorial team, especially Brigitte Séroussi, Lina Soualmia, Adrien Ugon, and Martina Hutter, as well as the reviewers who contributed to the selection process of the best papers.

References

1. Stanfill MH, Williams M, Fenton SH, Jenders RA, Hersh WR. A systematic literature review of automated clinical coding and classification systems. *J Am Med Inform Assoc* 2010;17(6):646-51.
2. Stanfill MH, Marc D. Health information management: implications of artificial intelligence on healthcare data and information management. *Yearb Med Inform* 2019;56-64.
3. Uzuner O, Luo Y, Szolovits P. Evaluating the state-of-the-art in automatic de-identification. *J Am Med Inform Assoc* 2007;14(5):550-63.
4. Du L, Xia C, Deng Z, Lu G, Xia S, Ma J. A machine learning based approach to identify protected health information in Chinese clinical text. *Int J Med Inform* 2018;116:24-32.
5. Zhang Y, Wang X, Hou Z, Li J. Clinical named entity recognition from Chinese electronic health records via machine learning methods. *JMIR Med Inform* 2018;6(4):e50.
6. Lee W, Kim K, Lee EY, Choi J. Conditional random fields for clinical named entity recognition: A comparative study using Korean clinical texts. *Comput Biol Med* 2018;101:7-14.
7. Bodenreider O, Cornet R, Vreeman DJ. Recent developments in clinical terminologies - SNOMED CT, LOINC, and RxNorm. *Yearb Med Inform* 2018;27(1):129-39.
8. Parr SK, Shotwell MS, Jeffery AD, Lasko TA, Matheny ME. Automated mapping of laboratory tests to LOINC codes using noisy labels in a national electronic health record system database. *J Am Med Inform Assoc* 2018;25(10):1292-300.
9. Cui L, Xie X, Shen Z. Prediction task guided representation learning of medical codes in EHR. *J Biomed Inform* 2018;84:1-10.
10. Atutxa A, Perez A, Casillas A. Machine learning approaches on diagnostic term encoding with the ICD for clinical documentation. *IEEE J Biomed Health Inform* 2018;22(4):1323-9.
11. Perez A, Atutxa A, Casillas A, Gojenola K, Sellart A. Inferred joint multigram models for medical term normalization according to ICD. *Int J Med Inform* 2018;110:111-7.
12. Qiu JX, Yoon H-J, Fearn PA, Tourassi GD. Deep learning for automated extraction of primary sites from cancer pathology reports. *IEEE J Biomed Health Inform* 2018;22(1):244-51.
13. Tang R, Ouyang L, Li C, He Y, Griffin M, Taghian A, et al. Machine learning to parse breast pathology reports in Chinese. *Breast Cancer Res Treat* 2018;169(2):243-50.
14. Miao S, Xu T, Wu Y, Xie H, Wang J, Jing S, et al. Extraction of BI-RADS findings from breast ultrasound reports in Chinese using deep learning approaches. *Int J Med Inform* 2018;119:17-21.
15. Linkov F, Silverstein JC, Davis M, Crocker B, Hao D, Schneider A, et al. Integration of cancer registry data into the text information extraction system: leveraging the structured data import tool. *J Pathol Inform* 2018;9:47.
16. Li F, Liu W, Yu H. Extraction of information related to adverse drug events from electronic health record notes: design of an end-to-end model based on deep learning. *JMIR Med Inform* 2018;6(4):e12159.
17. Chu J, Dong W, He K, Duan H, Huang Z. Using neural attention networks to detect adverse medical events from electronic health records. *J Biomed Inform* 2018;87:118-30.
18. Afzal N, Mallipreddi VP, Sohn S, Liu H, Chaudhry R, Scott CG, et al. Natural language processing of clinical notes for identification of critical limb ischemia. *Int J Med Inform* 2018;111:83-9.
19. Leroy G, Gu Y, Pettygrove S, Galindo MK, Arora A, Kurzius-Spencer M. Automated extraction of diagnostic criteria from electronic health records for autism spectrum disorders: development, evaluation, and application. *J Med Internet Res Journal* 2018;20(11):e10497.
20. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, Fifth Edition. Arlington, VA: American Psychiatric Association; 2013.

Correspondence to:
Meryl Bloomrosen
E-mail: meryl_bloomrosen@premierinc.com

Appendix: Content Summaries of Best Papers for the Health Information Management Section of the 2019 IMIA Yearbook

Atutxa A, Pérez A, Casillas A

Machine Learning Approaches on diagnostic term encoding with the ICD for clinical documentation

IEEE J Biomed Health Inform 2018;22(4):1323-9

This study focuses on data mining applied to unstructured clinical text in electronic health records (EHRs). The authors tried to improve standard machine learning techniques. They believe that clinical text mining can efficiently leverage the encoding process and they sought to develop computer-assisted classification tools and applications to help coding experts. The paper focuses on developing automatic techniques to encode diagnostic terms (DTs), focusing on Spanish language EHRs and publicly available resources for Spanish clinical text processing and mining.

In the study, records were encoded manually by experts using the International Classification of Diseases (ICD). The dataset consisted of spontaneous DTs extracted from over 9,000 EHRs, with 1,500 possible ICD codes. The proposed system was able to select the correct ICD code with 92% precision for the main disease (primary class) and 88% precision for the main disease together with the non-essential modifiers (fully specified class). The authors note that the methodology is simple and portable with potential applicability for documentation and pharmaco-surveillance. In a pilot study using a small sample of records, experts from public hospitals reported an accuracy of 91.2%. The authors have made the software publicly available so that the techniques and approach can be used by a broader audience of both clinicians and researchers.

Cui L, Xie X, Shen Z

Prediction task guided representation

learning of medical codes in EHR

J Biomed Inform 2018;84:1-10

The authors review applications using machine learning models for predictive analytics in electronic health records (EHRs). Machine learning has been used to improve the quality and efficiency of services. Developing machine learning models requires converting medical codes representing diagnoses and procedures to feature vectors. The authors recognize the importance of vector representations on the performance of machine learning models. They sought to address shortcomings of previous efforts using representation learning methods from Natural Language Processing (NLP) to learn vector representations of medical codes. As stated by the authors, “the objective of the study was to develop a representation learning model which can learn vector representations of medical codes that have strong predictive capability for various prediction tasks and required relatively small amounts of training data”. The researchers used a dataset that contained 750,000 cases and represented three years of records from five hospitals in China. Researchers developed a new method that they called “Prediction Task Guided Health Record Aggregation (PTGHRA)” which aggregates health records guided by prediction tasks, to construct a training corpus for various representation learning models. PTGHRA uses representation learning methods to map medical codes to continuous vectors and combined medical code vectors with other information in health record to form feature vectors for prediction tasks. Authors focused on the prediction of cost and length of stay (LoS). Compared with unsupervised approaches, representation learning models integrated with PTGHRA yielded a significant improvement in predictive capability of generated medical code vectors. For training set sizes smaller than 20,000 records, PTGHRA achieved up to 32% accuracy improvement.

Li F, Liu W, Yu H

Extraction of information related to adverse drug events from electronic health record notes: design of an end-to-end model

based on deep learning

JMIR Med Inform 2018;6(4):e12159

This paper focuses on how pharmacovigilance and drug-safety surveillance are crucial for monitoring adverse drug events (ADEs) and notes challenges (such as under-reporting) with several existing ADE-reporting systems. The main purpose of the study was to develop a deep learning model focusing on identifying ADEs, medications, and indications. A secondary purpose was to improve the deep learning model. The authors used the Medication, Indication, and Adverse Drug Events (MADE) 1.0 challenge to develop both training and testing datasets. MADE contains data (1,089 EHR notes) from cancer patients and includes nine entity types including Medication, Indication, and ADE. It also includes seven types of relations between these entities. The training data included 876 and the testing dataset used for the remaining 213 notes. To extract information from the dataset, the authors used a deep-learning model applying bidirectional long short-term memory (BiLSTM) conditional random field network to recognize entities and a BiLSTM-Attention network to extract relations. They enhanced their deep learning model with three multitask learning (MTL) methods (hard parameter sharing, parameter regularization, and task relation learning). The authors used the results of the second step of the process of extracting ADE information (relation extraction) to compare all models. They used micro-averaged precision, recall, and F1 as evaluation metrics. The authors compared their model with the top three systems in the MADE 1.0 challenge. Their model achieved state-of-the-art results (F1=65.9%). The model using hard parameter sharing further improved the F1 by 0.8%, boosting the F1 to 66.7%. The authors concluded that the performance of ADE-related information extraction can be improved by employing deep learning and MTL models. They also felt that the methods, data, and other factors influence the effectiveness of MTL models. The authors believe that this study could be useful for further natural language processing and machine learning research for detection of adverse drug events in clinical notes. Their annotated Dataset (Medication, Indication,

and Adverse Drug Events (MADE)) -- will be publicly available to support research on extraction of ADE-related information.

Qiu JX, Yoon H-J, Fearn PA, Tourassi GD

Deep learning for automated extraction of primary sites from cancer pathology reports

IEEE J Biomed Health Inform

2018;22(1):244-51

The paper focuses on using data extracted from pathology reports to populate cancer registries. The authors explore using a machine learning approach to make the labor-intensive manual process of information extraction and coding easier. In

this study, researchers investigated deep learning and a convolutional neural network (CNN) for extracting ICD-O-3 (International Classification of Diseases for Oncology) topographic codes from breast and lung cancer pathology reports. They compared a CNN with a term frequency vector approach, using 942 de-identified pathology reports matched to 12 ICD-O-3 topography codes corresponding to seven breast and five lung primary sites. They compared the results to reviews by human (cancer registry) subject matter experts. Pathology reports were provided from five different Surveillance, Epidemiology and End Results (SEER) cancer registries. Cancer registry experts manually annotated the pa-

thology reports and their annotations served as the gold standard. Researchers observed that when class labels were well populated, deep learning models (e.g., CNN) outperformed the conventional approaches (term frequency vector approaches) in the studies looking at class prevalence (micro- and macro-F score increases of up to 0.132 and 0.226). The increase in performance of deep learning methods from transfer learning was less strong and depended on the CNN method and cancer site. The authors believe that their results demonstrate the potential of deep learning for automated abstraction of relevant information for cancer registries from pathology reports.