Extraction of Active Medications and Adherence Using Natural Language Processing for Glaucoma Patients

Wei-Chun Lin, MD, MS[1], Jimmy S. Chen, MD[2], Joel Kaluzny, MD[3], Aiyin Chen, MD[3], Michael F. Chiang, MD[4], Michelle R. Hribar, PhD[1]
[1]Medical Informatics & Clinical Epidemiology, [2]School of Medicine, [3]Ophthalmology Oregon Health & Science University, Portland, OR, [4]National Eye Institute, Bethesda, MD

Abstract

Accuracy of medication data in electronic health records (EHRs) is crucial for patient care and research, but many studies have shown that medication lists frequently contain errors. In contrast, physicians often pay more attention to the clinical notes and record medication information in them. The medication information in notes may be used for medication reconciliation to improve the medication lists' accuracy. However, accurately extracting patient's current medications from free-text narratives is challenging. In this study, we first explored the discrepancies between medication documentation in medication lists and progress notes for glaucoma patients by manually reviewing patients' charts. Next, we developed and validated a named entity recognition model to identify current medication and adherence from progress notes. Lastly, a prototype tool for medication reconciliation using the developed model was demonstrated. In the future, the model has the potential to be incorporated into the EHR system to help with real-time medication reconciliation.

Introduction

The rapid adoption of electronic health records (EHRs) has generated large-scale clinical data that has been re-used for many purposes, including patient phenotyping,[1] pharmacovigilance,[2, 3] comparative effectiveness research,[4] clinical decision support,[5, 6] and quality improvement and research.[7] Although secondary use of EHR shows many benefits such as improved healthcare quality, reduced healthcare costs, and effective clinical research,[8, 9] there are many challenges that still need to be addressed. One of the biggest challenges is the accuracy and completeness of EHR data, specifically medication information.[10]

The accuracy of medication data is crucial for patient safety, quality of care, and clinical research. Inaccurate or incomplete medication records can lead to polypharmacy, adverse medication interactions, and decreased data reliability in research.[11] The medication list is a structured record of a patient's medication data which is populated automatically by electronically prescribed medications or manually through medication reconciliation.[12] However, the EHR system may not always capture medication data correctly or prevent errors in the medication list.[13] Previous studies have shown that medication lists frequently contain errors, including duplicated documentation of medications, outdated discontinued prescriptions in the medication list, and missing medications prescribed elsewhere.[12, 14-18] In addition, prior studies show that physicians direct very little attention to EHR medication lists, and instead spend most time reviewing the impression and plan section.[19, 20] It seems reasonable to expect that medications recorded in narrative notes are more reliable and can be helpful with medication reconciliation. Medication reconciliation is a process to create and maintain patients' most current and accurate list of medications.[21, 22]

However, manual reviewing progress notes for medication data extraction in EHR is time-consuming and labor-intensive. Natural language processing (NLP) is a promising strategy for capturing medication information from the free-text progress note. With advancements in machine learning and the large text corpora available in EHR, NLP has been successfully used to process free-text EHR data, for deep contextualized word representations,[23] information extraction,[24] and semantic analysis.[25] Named entity recognition (NER) is a sub-task of information extraction, which seeks to identify words or phrases into pre-defined categories with specific labels.[26] Over the past years, the NER technique has been applied to extract medication information, such as drug names, frequency, dosage, adverse drug events, adherence, etc., from free-text documents.[27-31] For example, a conditional random field (CRF) model was used to develop a NER model to detect medication attributes and adverse drug events.[28] Also, bidirectional long short-term memory (LSTM) model was used for named entity recognizing for medication information.[27, 29] More recently, pre-trained deep learning models were widely used for biomedical information extraction.[31, 32] However, to our knowledge, there has not been a well-developed NLP tool to identify a list of *current medications* for a specific disease such as glaucoma and help with medication reconciliation.

The purpose of this study is to develop a NER model for extracting patients' current ophthalmologic medication and adherence from free-text notes for glaucoma patients. Glaucoma is characterized by progressive degeneration of the optic nerve and irreversible visual field loss, and it is the leading cause of irreversible blindness worldwide.[33] The majority of glaucoma patients are treated using medical therapy, and the accuracy of medication documentation is crucial in glaucoma management [34] However, the accuracy of glaucoma medication documentation is unclear. In addition, glaucoma patients' medication non-adherence rate has been reported to vary from 24% to 59%.[35, 36] Therefore, a reliable method to assess glaucoma patients' current ophthalmologic medication and adherence is needed.[37] Finally, the reliability of medication data is important for glaucoma research, such as prediction models for disease progression. In this study, we first manually reviewed patient charts for discrepancies in medication documentation between medication lists and progress notes. Next, we trained and tested a NER model for extracting current medication from progress notes and evaluated its accuracy. Finally, we demonstrated an approach for medication reconciliation using the NER model on small sample progress notes.

Methods

This study was approved by the Institutional Review Board at Oregon Health and Science University (OHSU). OHSU is a large academic medical center in Portland, Oregon. This study was conducted at Casey Eye Institute, OHSU's ophthalmology department serving all major ophthalmology subspecialties. The department performs over 130,000 outpatient examinations annually and is a major referral center in the Pacific Northwest and nationally. In 2006, OHSU implemented an institution-wide EHR (EpicCare; Epic Systems, Verona, WI) to handle all ambulatory practice management, clinical documentation, order entry, medication prescribing, and billing.

The study contains three phases (1) Explore medication discrepancies between the medication list and the progress note for glaucoma by manually reviewing charts; (2) Develop a NER model to extract patients' current ophthalmologic medication and medication adherence from progress notes for glaucoma patients and (3) Apply the NER model to perform medication reconciliation.

1. Manual Chart Review of Medication Lists and Progress Notes

Progress notes and medication list data from EHR were extracted for 150 randomly selected Casey Eye Institute patients with encounter ICD10 diagnosis codes related to glaucoma from January 23, 2019, to September 28, 2020. The patient's most recent office visit notes were manually reviewed by three independent reviewers. The medications recorded in the narrative notes were abstracted and compared to the EHR medication list at the time of visit. All ophthalmologic medications and over-the-counter (OTC) medications (e.g., artificial tears) were collected. All medications listed in the notes but not on the medication list or vice versa were labeled. Cross-validation among the three reviewers was conducted by using a subset of 20 encounter notes (96.4% agreement).

2. NER Model for Extracting All Ophthalmic Medications

We sampled a dataset with 507 progress notes from office visits at the Casey Eye Institute from January 01, 2019, to December 31, 2019, with encounter ICD10 codes associated with glaucoma. The dataset was constructed by random stratified sampling from all ophthalmology visits according to the department and primary provider name. The documents were manually annotated for nine categories: Drug Name, Route, Frequency, Dosage, Strength, Duration, Adverse Drug Event (ADE), Adherence, and Current Medication Use. All medication names, including generic names, brand names, and abbreviations, were sourced from publicly available online resources and glaucoma specialists. An open-source tool (Doccano; Open source: Doccano; 2018) was used to annotate the documents.[38] Due to the limited number of ADE entities, we discarded this category and kept the other eight entities. **Figure 1** displays an example of the annotation. The annotated dataset was randomly split into 75% for training and 25% for testing. A 10% randomly sampled subset of documents from the training data was used as a validation set for turning the hyperparameters. **Table 1** presents the description of the datasets and annotation statistics.

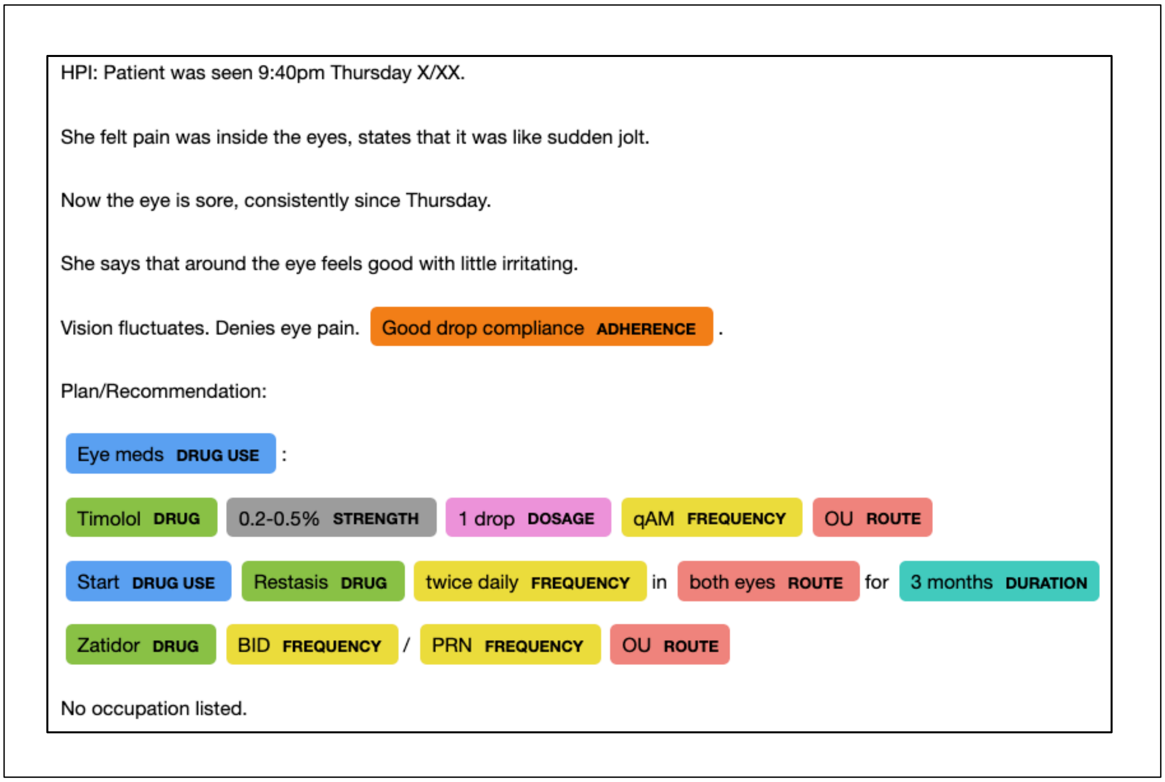

Figure 1. Example of note annotation by an open-source tool. Medication drug name, strength, dosage, frequency, route, duration, current use, and adherence are identified.

We used named entity recognition, a sub-type information extraction technique, to extract medication information and adherence from clinical notes. The NER model was developed in Python 3.7.6 using the spaCy library.[39,40] The spaCy library is a free open-source library for NLP. The architecture of spaCy's NER model is based on convolutional neural networks which uses a word embedding strategy using sub-word features and "Bloom" embeddings.[41,42] In this study, the training task contains 200 epochs with experiments with multiple hyperparameter settings. Different learning rates (initial at 1e−2, 1e−3, 2e−3, 1e−4, 2e−4) were tested and adjusted by two optimizers: Adaptive Moment Estimation (Adam) and stochastic gradient descent. We use a decaying dropout rate (0.5 - 0.35; 1e-3) to avoid overfitting. Also, we experimented with different batch compounding sizes and regularization schemes to optimize the model. The results of the NER model's extraction for the test set were determined by comparing the manually annotated and the NER model's extracted entities. The model performance was evaluated by using F1 score, precision, recall, and the micro-averaged score, which aggregates the contributions of all categories to calculate the average metrics.[43]

Table 1. Distribution of annotated entities and number of progress notes in training and testing datasets.

Named Entities	Train	Test	Total
Drug	2029	505	2534
Frequency	1722	411	2133
Route	1666	371	2037
Dosage	201	40	241
Duration	35	15	50
Strength	168	31	199
Adherence	132	48	180
Current Medication Use	725	185	910
Number of Notes	381	126	507

3. Medication Reconciliation Using NER Model for Current Medications

Finally, we developed a prototype medication reconciliation tool using the optimized NER model. For this purpose, we are focusing only on medications the patient is currently using as documented in progress notes. **Figure 2** demonstrates an example of medication reconciliation using our prototype tool. First, our NER model extracted the patient's medications and "Drug Use" label from the150 sample progress notes which were manually reviewed in phase 1. The "Drug Use" label identified which medications that the patient was currently taking. Next, the current medications were standardized based on RxNorm Ingredient (IN).[44] Finally, the standardized medications were compared to the manually identified medications from phase 1. Both ophthalmologic medications and over-the-counter (OTC) medications (e.g., artificial tears) were included. All medications listed in the notes but not on the medication list or vice versa were flagged.

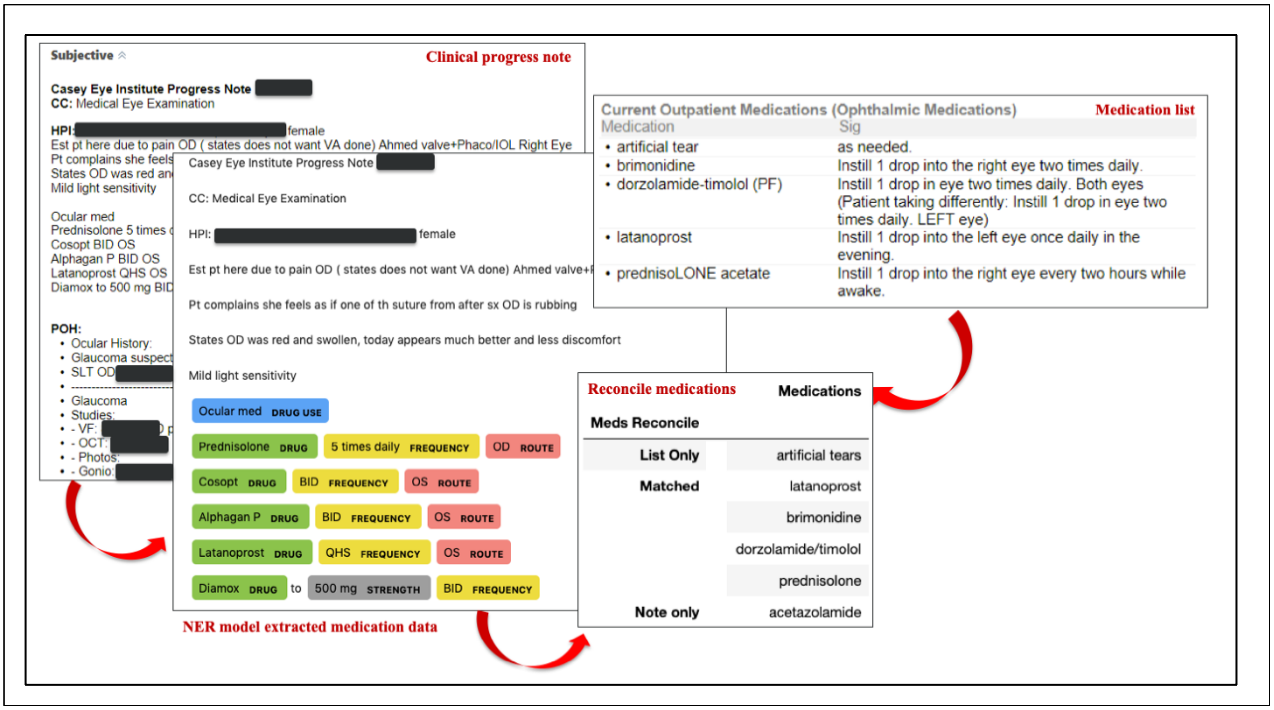

Figure 2. Example of medication reconciliation using the developed NER model

Results

1. Manual Chart Review of Medication Lists and Progress Notes

The randomly sampled 150 patients' notes and medication lists contained a total of 450 medications, including glaucoma eye drops, mydriasis eye drops, antimicrobials, corticosteroids, and OTC medications. Prescription medications were most common (n = 355; 79%), followed by OTC medications (n = 95; 22%). Around 57% of patients had at least one medication mismatch for all categories in their records. However, only 36% of patients had at least one medication mismatch for prescription medications (**Figure 3**). Nearly 66% of medications (n = 298) could be reconciled between the progress notes and medication list. Around 34% (n = 152) of medications are mismatched for various reasons, including medications prescribed by clinicians from different institutions, medications with duplicated prescriptions, medications that were prescribed and entered in the medication list but not recorded in the progress note, and old medications that were not discontinued in the medication list. **Figure 4** displays the distribution of medication mismatches among the two categories in the EHR by location. The most frequent mismatch was found with prescription medications (55%) followed by the OTC medications (45%). The OTC medications were more commonly recorded in the progress notes but not entered into the medication list. In contrast, mismatched prescription medication more often appeared in the medication list but not in the progress notes.

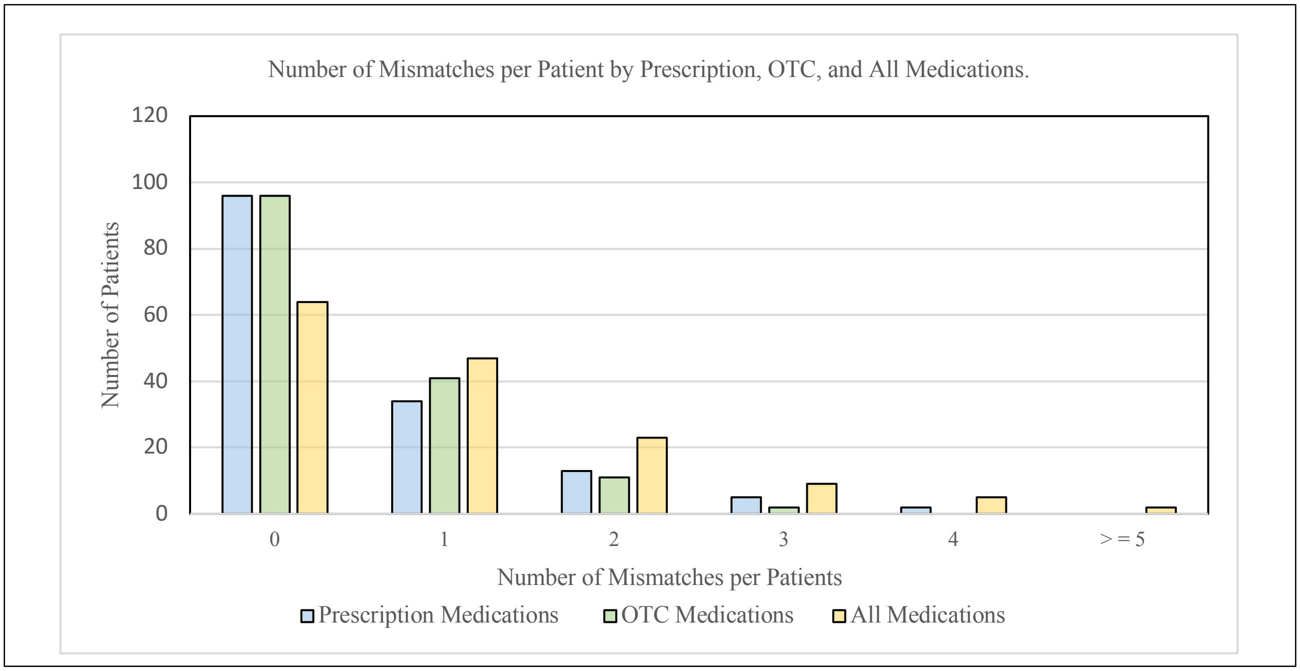

Figure 3. Medication documentation mismatches were stratified based on the number of mismatches that occurred per patient for prescription (blue), OTC (green), and all medications (yellow).

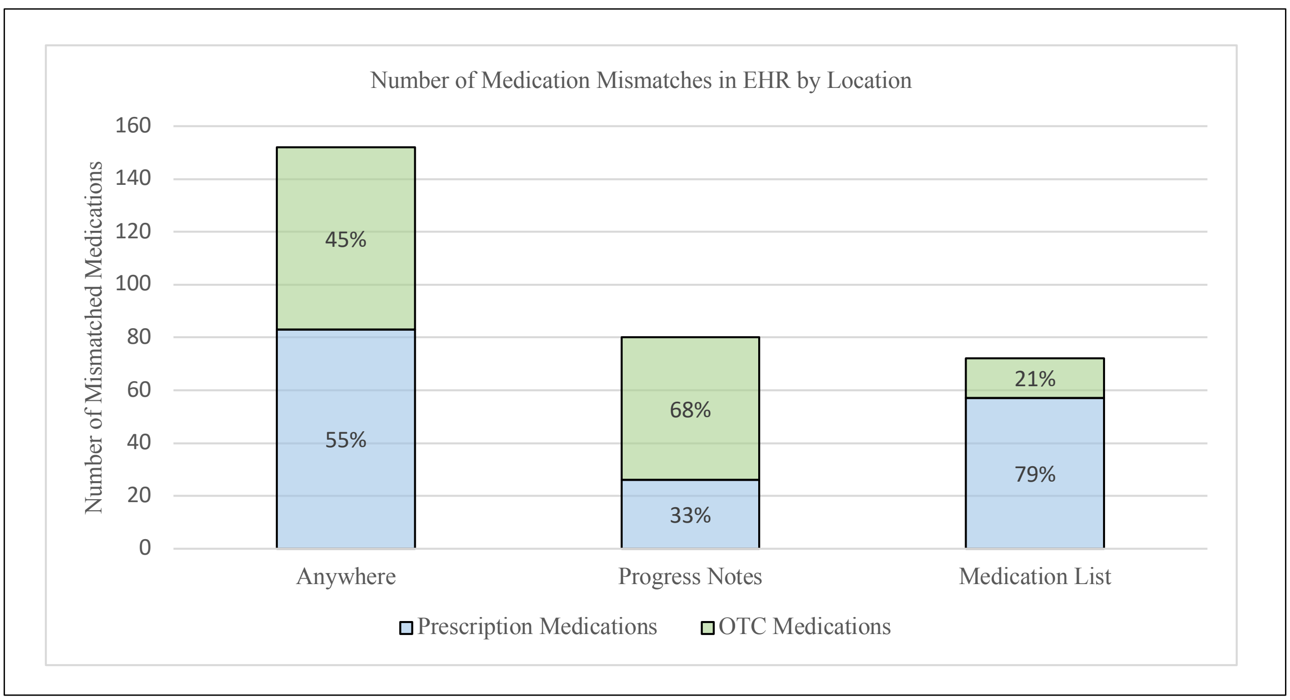

Figure 4. Summary of medication mismatches across 150 patients.

2. NER Model for Extracting Current Ophthalmic Medication

The custom NER model was trained with 381 progress note documents that were manually annotated with eight named entities and then tested on 126 progress notes. **Table 2** presents the overall micro-averaged and per-entity performance for the optimal NER model on test data (126 progress notes). The overall performance of the NER model across all categories was F1 score = 0.955, Precision = 0.951, and Recall = 0.957. Higher performance was observed on medication-related entities: Drug, Name, Route, Frequency, Dosage, and Strength, compared to patient's behavior-related entities: Adherence and Current Medication Use. An error analysis was performed for false negative and positive on Drug Name, Adherence, and Current Medication Use to recognize the source of error predictions. Several causes of errors were identified, such as different wordings for medication adherence, mislabeled current medication use and drug name due to similar sentence structure, eye exams or warm compress mislabeled as drug name, and misclassification when entity information was contained in a short sentence. (**Table 3**).

Table 2. The results of the NER model on the test dataset

Entities	Performance on Test Data		
	Precision	**Recall**	**F1-Score**
Drug	0.971	0.971	0.971
Frequency	0.972	0.969	0.970
Route	0.948	0.986	0.966
Dosage	0.987	0.998	0.991
Duration	1.000	0.600	0.749
Strength	0.969	0.997	0.982
Adherence	0.803	0.758	0.779
Current Medication Use	0.899	0.919	0.909
Average (micro)	0.951	0.957	0.955

Table 3. Error analysis from NER predictions related to Drug Name, Current Medication Use, and Adherence labels

Error category	Example	Explanation
Mislabeled Current Medication Use	**"Urgent add on** - Last seen Dr. X on X/X/XXXX"	Unexplained error, "Urgent add on" was labeled as Current Medication Use
	"**encouraged** PFATs at least BID OU - discussed to space at least 5 mins from glaucoma drops"	"Encouraged" was mislabeled as Current Medication Use due to the similar sentence structure
Mislabeled Adherence	"- History *inconsistent drop adherence*"	There are many different wordings for medication adherence, and Adherence label was not assigned
	"No eye pain/discomfort but patient admits to *forgetting his drops frequently*."	
Mislabeled Drug Name	"Cont **warm compresses** BID ou"	"warm compresses" was mislabeled as Drug Name due to a similar sentence structure
	"Vision has been good. Just **using OTC readers**."	"Using" was mislabeled as Current Medication Use, and OTC readers was mislabeled as Drug Name due to the similar sentence structure

3. Medication Reconciliation Using the NER Model

The prototype medication reconciliation tool identified 408 current medications from the 150 progress notes that were manually reviewed in phase 1. After standardizing the medications to RxNorm, 14 medications were removed for a final list of 394 medications. Among the 394 medications, there were 379 medications matched with the manually abstracted current medications. The prototype tool achieved a good performance of F1 score = 0.969, Precision = 0.959, and Recall = 0.979.

Discussion

In this study, we explored medication discrepancies in the EHR data and evaluated the performance of a custom NER model's applicability to extract current medication for glaucoma patients. We also used the developed NER model in a proof-of-concept application to perform medication reconciliation in a subset of our patients. The key findings from our study were (1) Medication discrepancies in patient charts were found to be present in a large proportion of office visits; (2) The custom NER model can accurately extract current medication and adherence for glaucoma patients; (3) The NER model can be used to reconcile the medication documentation.

The first key finding is that medication discrepancies were found to be present in a large proportion of office visits. Our study shows that approximately twenty percent of medications prescribed to glaucoma patients had at least one discrepancy between the medication list and the progress note. Overall, more than one-third of patients in this study had at least one medication mismatch between both data sources. These inconsistencies in the EHR medication records may increase the risk of medication errors[45] and affect the reliability of research that relies on this data. These findings are similar to other studies, including a study for microbial keratitis demonstrating 76.9% of medication agreement between progress notes and medication lists[12] and another study for inflammatory bowel disease reporting 78.6% of medications agreement between clinical narrative and medication list.[15] The findings from these studies indicate that the accuracy of the medication list is a common problem. An accurate tool for medication reconciliation of medication lists and further qualitative studies to understand the causes of medication data discrepancies is needed.

The second key finding is that our NER model can accurately identify current medication and adherence from progress notes from outpatient glaucoma visits. In our study, the model reached a micro-averaged F1 score of 0.955 across all categories. The NER model was developed to recognize eight categories from free-text progress notes, including drug names (including generic, brand, and abbreviation names), the route of administration, prescription frequency, the dosage of the drugs, drug strength, duration, medication adherence, and current medication use. The NER model could accurately identify medication-related entities (except duration) but showed lower performance on patient behavior-related entities, such as adherence and current medication use. The difference could be ascribed to the limited number of training cases and the higher variety of wordings. As shown in **Table 1**, there are only 35 annotated duration entities and 132 annotated adherence entities in the training data. In addition, the words and phrases to indicate adherence and current medication use are various, and some of these phrases are located in different sentences than the medications. Nevertheless, the most common error of drug name identification is mislabeling other terms such as "warm compress" or "OTC readers" as a drug name due to similar sentence structure. For example, "warm compress left eye PRN" or "Vision has been good. Just using OTC readers." In these cases, these mislabeled drugs will easily be filtered out of the results in practice during the conversion to RxNorm names.

Finally, our NER model can be used to reconcile medication documentation. As shown in the phase one study, we can manually abstract the medication records from their progress notes to compare with their medication list. Similarly, the NER model was able to recognize common medications as well as identify text related to current medication use. This is the first study, to our knowledge, to develop NLP models to recognize *current medication* use from free-text progress notes. With the ability to identify the current medication use, we are able to capture the whole picture of current medications for the target patients and reconcile it with their medication list. As previously mentioned, the medication reconciliation between progress notes and medication lists was only reported from 76.9% to 79.6% for three different diseases, including microbial keratitis, inflammatory bowel disease, and glaucoma.[12, 15] And more than one-third of patients had at least one discrepancy for ophthalmic prescription medications.[12] In our study, the NLP tool can correctly identify current medications for glaucoma patients on 150 sample progress notes (F1 score = 0.969). **Figure 2** displays an example of medication reconciliation using the NLP tool. In this prototype tool, we focused on reconciling the drug names since physicians did not always record the other attributes, such as route, frequency, and dosage along with the medications. In future work, we plan to extend this medication reconciliation method to use the information from both narrative progress notes and medication lists to construct a current medication list for glaucoma patients.

Our study has limitations future work may address. First, some of the entities are naturally less frequently recorded in the progress notes that affect the performance of the NER model. For example, text related to drug duration appeared much less frequently than other entities, such as drug name, route, and frequency. Thus, it is challenging to train the model correctly recognize these entities. A similar finding was reported in another study.[31] Second, the model was trained on a set of notes for glaucoma patients from a single institution; it is unclear if the model can be generalizable to other subspecialties within ophthalmology or other healthcare systems. Finally, the application of the custom NER model for medication reconciliation is a proof of concept. We conducted the test of medication reconciliation using the NER model on a limited number of samples. Our intention is to extend and replicate these study methods to different specialties and institutions to increase the generalizability of our model. In the future, the custom model could be incorporated into the EHR system to help with medication reconciliation.

Conclusion

Discrepancies in medication documented in the medication and in progress notes were observed in more than one-third of encounters for glaucoma patients. Inaccurate medication lists in the EHR may affect the reliability of the research or clinical decision support using this data. Since physicians often record current medication information in the progress notes this data could be used for medication reconciliation. In this study, we developed an NLP model to accurately identify current medication information from free-text EHR data that can be applied to perform automated medication reconciliation; the performance of the model is similar to the best performing published NLP models for medication extraction studies.[27-31, 46] This has implications in improving the data quality and usefulness for medication data in both research and clinical care.

Acknowledgments

Supported by grants T15LM007088, 1R21EY031443-01, and P30EY0105072 from the National Institutes of Health (Bethesda, MD) and unrestricted departmental support from Research to Prevent Blindness (New York, NY). MFC was a consultant for Novartis (Basel, Switzerland) and previously an equity owner in InTeleretina LLC (Honolulu, HI), and received research support from Genentech (San Francisco, CA) and the National Science Foundation (Alexandria, VA).

References

1. Chute CG, Ullman-Cullere M, Wood GM, Lin SM, He M, Pathak J. Some experiences and opportunities for big data in translational research. *Genet Med*. 2013;15(10):802-809.
2. LePendu P, Iyer SV, Fairon C, Shah NH. Annotation analysis for testing drug safety signals using unstructured clinical notes. *J Biomed Semantics*. 2012:1-12.
3. Warrer P, Hansen EH, Juhl-Jensen L, Aagaard L. Using text-mining techniques in electronic patient records to identify ADRs from medicine use. *Br J Clin Pharmacol*. 2012;73(5):674-684.
4. Forrest CB, Margolis PA, Bailey LC, et al. PEDSnet: a national pediatric learning health system. *J Am Med Inform Assoc*. 2014;21(4):602-606.
5. Rothman B, Leonard JC, Vigoda MM. Future of electronic health records: implications for decision support. *Mt Sinai J Med*. Nov-Dec 2012;79(6):757-768. doi:10.1002/msj.21351
6. Wagholikar KB, Sundararajan V, Deshpande AW. Modeling paradigms for medical diagnostic decision support: a survey and future directions. *J Med Syst*. Oct 2012;36(5):3029-3049. doi:10.1007/s10916-011-9780-4
7. Rowley J. The wisdom hierarchy: representations of the DIKW hierarchy. *J Inform Sci* 2007;33(2):163-180.
8. Meystre SM, Lovis C, Bürkle T, Tognola G, Budrionis A, Lehmann CU. Clinical data reuse or secondary use: current status and potential future progress. *Yearb Med Inform*. 2017;26(1):38.
9. Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol*. Jan 2017;106(1):1-9. doi:10.1007/s00392-016-1025-6
10. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J Biomed Inform*. 2013;46(5):830-836.

11. FitzGerald RJ. Medication errors: the importance of an accurate drug history. *Br J Clin Pharmacol*. Jun 2009;67(6):671-675. doi:10.1111/j.1365-2125.2009.03424.x
12. Ashfaq HA, Lester CA, Ballouz D, Errickson J, Woodward MA. Medication accuracy in electronic health records for microbial keratitis. *JAMA Ophthalmol*. May 30 2019;137(8):929-931. doi:10.1001/jamaophthalmol.2019.1444
13. Mekonnen AB, McLachlan AJ, Brien JaE. Pharmacy‐led medication reconciliation programmes at hospital transitions: a systematic review and meta‐analysis. *J Clin Pharm Ther*. 2016;41(2):128-144.
14. Woodward MA, Ballouz D, Errickson J, Maganti N, Tuohy M. Medication burden for patients with bacterial keratitis. *Invest Ophthalmol Vis Sci*. 2019;60(9):841-841.
15. Walsh KE, Marsolo KA, Davis C, et al. Accuracy of the medication list in the electronic health record-implications for care, research, and improvement. *J Am Med Inform Assoc*. Jul 1 2018;25(7):909-912. doi:10.1093/jamia/ocy027
16. Caglar S, Henneman PL, Blank FS, Smithline HA, Henneman EA. Emergency department medication lists are not accurate. *J Emerg Med*. 2011;40(6):613-616.
17. Orrico KB. Sources and types of discrepancies between electronic medical records and actual outpatient medication use. *J Manag Care Pharm*. 2008;14(7):626-631.
18. Chan KS, Fowles JB, Weiner JP. Electronic health records and the reliability and validity of quality measures: a review of the literature. *Med Care Res Rev*. Oct 2010;67(5):503-527. doi:10.1177/1077558709359007
19. Brown P, Marquard J, Amster B, et al. What do physicians read (and ignore) in electronic progress notes? *Appl Clin Inform*. 2014;5(2):430.
20. Rosenbloom ST, Denny JC, Xu H, Lorenzi N, Stead WW, Johnson KB. Data from clinical notes: a perspective on the tension between structure and flexible documentation. *J Am Med Inform Assoc*. 2011;18(2):181-186.
21. Kwan JL, Lo L, Sampson M, Shojania KG. Medication reconciliation during transitions of care as a patient safety strategy: a systematic review. *Annals of internal medicine*. 2013;158(5_Part_2):397-403.
22. Moghli MA, Farha RA, Hammour KA. Medication discrepancies in hospitalized cancer patients: Do we need medication reconciliation? *J Oncol Pharm Pract*. 2021;27(5):1139-1146.
23. Peters ME, Neumann M, Iyyer M, et al. Deep contextualized word representations. *arXiv preprint arXiv:180205365*. 2018;
24. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. *J Biomed Inform*. Jan 2018;77:34-49. doi:10.1016/j.jbi.2017.11.011
25. del Carmen Legaz-García M, Martínez-Costa C, Menárguez-Tortosa M, Fernández-Breis JT. A semantic web based framework for the interoperability and exploitation of clinical models and EHR data. *Knowl Based Syst*. 2016;105:175-189.
26. Mohit B. Named entity recognition. *Natural language processing of semitic languages*. Springer; 2014:221-245.
27. Wunnava S, Qin X, Kakar T, Sen C, Rundensteiner EA, Kong X. Adverse drug event detection from electronic health records using hierarchical recurrent neural networks with dual-level embedding. *Drug saf*. 2019;42(1):113-122.
28. Chapman AB, Peterson KS, Alba PR, DuVall SL, Patterson OV. Detecting adverse drug events with rapidly trained classification models. *Drug saf*. Jan 2019;42(1):147-156. doi:10.1007/s40264-018-0763-y
29. Magge A, Scotch M, Gonzalez-Hernandez G. Clinical NER and relation extraction using bi-char-LSTMs and random forest classifiers. PMLR; 2018:25-30.
30. Tarcar AK, Tiwari A, Dhaimodker VN, Rebelo P, Desai R, Rao D. Healthcare NER models using language model pretraining. *arXiv preprint arXiv:191011241*. 2019;
31. Kormilitzin A, Vaci N, Liu Q, Nevado-Holgado A. Med7: a transferable clinical natural language processing model for electronic health records. *arXiv preprint arXiv:200301271*. 2020;
32. Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*. Feb 15 2020;36(4):1234-1240. doi:10.1093/bioinformatics/btz682
33. Flaxman SR, Bourne RR, Resnikoff S, et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *Lancet Glob Health*. Dec 2017;5(12):e1221-e1234. doi:10.1016/S2214-109X(17)30393-5
34. Tsai JC. Medication adherence in glaucoma: approaches for optimizing patient compliance. *Curr Opin Ophthalmol*. Apr 2006;17(2):190-195. doi:10.1097/01.icu.0000193078.47616.aa
35. Gurwitz JH, Glynn RJ, Monane M, et al. Treatment for glaucoma: adherence by the elderly. *Am J Public Health*. May 1993;83(5):711-716. doi:10.2105/ajph.83.5.711

36. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. Aug 4 2005;353(5):487-497. doi:10.1056/NEJMra050100
37. Muir KW, Lee PP. Glaucoma medication adherence: room for improvement in both performance and measurement. *Arch Ophthal*. 2011;129(2):243-245.
38. Doccano: Open Source text annotation for machine learning practitioner [Internet]. Updated 2018. Accessed December 15, 2020. https://github.com/doccano/doccano
39. Honnibal M, Montani I. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. *To appear*. 2017;7(1):411-420.
40. Van Rossum G. Python Programming Language. 2007:36.
41. Serra J, Karatzoglou A. Getting deep recommenders fit: Bloom embeddings for sparse binary input/output networks. 2017:279-287.
42. Lample G, Ballesteros M, Subramanian S, Kawakami K, Dyer C. Neural architectures for named entity recognition. *arXiv preprint arXiv:160301360*. 2016;
43. Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. Springer; 2005:345-359.
44. Liu S, Ma W, Moore R, Ganesan V, Nelson S. RxNorm: prescription for electronic drug information exchange. *IT professional*. 2005;7(5):17-23.
45. Hellström LM, Bondesson Å, Höglund P, Eriksson T. Errors in medication history at hospital admission: prevalence and predicting factors. *BMC Clin Pharmacol*. 2012;12(1):1-9.
46. Ngo D-H, Metke-Jimenez A, Nguyen A. Knowledge-based feature engineering for detecting medication and adverse drug events from electronic health records. PMLR; 2018:31-38.