

# Natural Language Processing for Automated Annotation of Medication Mentions in Primary Care Visit Conversations

Craig H Ganoë, MS<sup>1</sup>, Weiyi Wu, MS<sup>1</sup>, Paul J Barr, MSc, PhD<sup>2</sup>, William Haslett, PhD<sup>1</sup>, Michelle D Dannenberg, MPH<sup>2</sup>, Kyra L Bonasia, PhD<sup>2</sup>, James C Finora<sup>2</sup>, Jesse A Schoonmaker, MD, MPH<sup>2</sup>, Wambui M Onsando, MD, MPH<sup>2</sup>, James Ryan, DO<sup>3</sup>, Glyn Elwyn, MSc, MD, PhD<sup>2</sup>, Martha L Bruce, MPH, PhD<sup>2</sup>, Amar K Das, MD, PhD<sup>2</sup>, Saeed Hassanpour, PhD<sup>1,4,5</sup>

<sup>1</sup> Biomedical Data Science Department, Geisel School of Medicine, Dartmouth College, Lebanon, NH 03756, USA

<sup>2</sup> The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine, Dartmouth College, Lebanon, NH 03756, USA

<sup>3</sup> Ryan Family Practice, Ludington, MI 49431, USA

<sup>4</sup> Computer Science Department, Dartmouth College, Hanover, NH 03755, USA

<sup>5</sup> Epidemiology Department, Geisel School of Medicine, Dartmouth College, Lebanon, NH 03756, USA

Corresponding Author: Saeed Hassanpour, PhD

Postal address: One Medical Center Drive, HB 7261, Lebanon, NH 03756, USA

Telephone: (603) 650-1983

Email: [Saeed.Hassanpour@dartmouth.edu](mailto:Saeed.Hassanpour@dartmouth.edu)

Key words: Clinic visit recording; Medication information extraction; Natural language processing

## ABSTRACT

**Objectives:** The objective of this study is to build and evaluate a natural language processing approach to identify medication mentions in primary care visit conversations between patients and physicians.

**Materials and Methods:** Eight clinicians contributed to a dataset of 85 clinic visit transcripts, and ten transcripts were randomly selected from this dataset as a development set. Our approach utilizes Apache cTAKES and Unified Medical Language System (UMLS) controlled vocabulary to generate a list of medication candidates in the transcribed text, and then performs multiple customized filters to exclude common false positives from this list while including some additional common mentions of the supplements and immunizations.

**Results:** Sixty-five transcripts with 1,121 medication mentions were randomly selected as an evaluation set. Our proposed method achieved an F-score of 85.0% for identifying the medication mentions in the test set, significantly outperforming existing medication information extraction systems for medical records with F-scores ranging from 42.9% to 68.9%.

**Discussion:** Our medication information extraction approach for primary care visit conversations showed promising results, extracting about 27% more medication mentions from our evaluation set while eliminating many false positives in comparison to existing baseline systems. We made our approach publicly available on the web as an open-source software.

**Conclusion:** Integration of our annotation system with clinical recording applications has the potential to improve patients' understanding and recall of key information from their clinic visits, and, in turn, to positively impact health outcomes.

## BACKGROUND AND SIGNIFICANCE

Forty to eighty percent of healthcare information is forgotten *immediately* by patients post visit.[1–4] Poor recall and understanding of medical concepts have been identified as significant barriers to self-management, a central component of the Chronic Care Model, resulting in poorer health outcomes.[5–7] These barriers are amplified in older adults with multimorbidity,[8–11] where reduced cognitive capacity,[12–14] low health literacy,[15,16] and complex treatment plans are common.[17–19] Older adults with multimorbidity account for 96% of Medicare expenditures, and in the absence of optimal self-management, they experience lower quality of life and greater functional decline.[10,11,20–26]

An after-visit summary, shared via a patient portal, is a common strategy to improve recall of visit information.[27–29] However, summaries impose a significant burden on clinicians who must document the entire visit in terms that are understandable to patients, with low health literacy being common.[30,31] Alternatively, audio recordings can provide a full account of the clinic visit and are an effective modality—71% of patients listen to recordings and 68% share their recording with a caregiver.[32] Clinic recordings improve patient understanding and recall of visit information, reduce anxiety, increase satisfaction, and improve treatment adherence.[32–38] As patient demand for recordings increases,[39,40] a growing number of clinics across the U.S. are offering audio recordings of clinic visits, and a recent survey reveals that almost a third of clinicians in the U.S. have shared a recording of a clinic visit with patients.[41]

Yet, unstructured clinic recordings may overwhelm patients.[39,42] Advances in data science methods, such as natural language processing (NLP), can be used to identify patterns in unstructured data and extract clinically meaningful information. These methods have been used to predict hospital readmissions,[43] future radiology utilizations,[44] and characterization of significance, change, and urgency of clinical findings in medical records.[45–49] As such, we plan to develop a recording system for patients that applies NLP methods to unstructured clinic visit recordings.[50]

In this paper, we describe an approach to extract mentions of medication names in transcripts of clinic visit audio recordings. Annotating mentions of medications discussed during a clinic visit recording can provide added value to the audio-recorded health information. We use NLP to highlight medication mentions in transcripts of clinic recordings. These annotations can be utilized to index the audio and aid visit recall by enabling key visit information

to be easily accessed. In addition, the indexed medical concepts can be linked to credible and trustworthy online resources. These resources would provide additional information about medications to aid in patient understanding. Such an approach could potentially increase patient self-management, and, when shared with caregivers, could increase their confidence in delivering care.

At the time of this work, no prior work focused on extracting medication information from clinic visit conversations and their transcriptions. There has been some work toward the extraction of medication names and also prescription-related attributes such as dosage and frequency from medical text (primarily clinical notes). These systems have mainly focused on the extraction of medication information from written clinical notes. In 2009, the Third i2b2 Shared-Task on Challenges in Natural Language Processing for Clinical Data Workshop focused on medication information extraction. The challenge was to extract and label medication-related terms (medication name, dosage, frequency, etc.) from discharge summaries.[51] Teams were given 696 summaries for development, and then 547 summaries were used for evaluation. Twenty teams submitted entries to the challenge, with the top result for annotating medication names being an F-score of 90.3% on the evaluation dataset, utilizing a combination of a rule-based approach with two machine learning models (conditional random field and support vector machine). This top approach also achieved an F-score of 90.81% on an internal test set of 30 clinical records when evaluated by the system's authors.[52]

Since the 2009 i2b2 challenge, additional work has been done to improve medication information extraction methods. Sohn et al. developed Medication Extraction and Normalization (MedXN) to extract medication information and map it to the most specific RxNorm concept possible.[53] This group reported an F-score of 97.5% for medication name on a test set of 26 clinical notes containing 397 medications. In 2014, MedEx, the system with the second-best results in the i2b2 challenge, was reimplemented using Unstructured Information Management Architecture (UIMA) to extract drug names and map them to both generalized and specific RxNorm concepts.[54] This system, named MedEx-UIMA, achieved an F-score of 97.5% for extracting and mapping to the most generalized concept and an F-score of 88.1% for mapping to the most specific concept, evaluating on a set of 125 discharge summaries from the original i2b2 challenge. The authors concluded that the new MedEx-UIMA implementation was consistent with and sometimes outperformed the original MedEx method. Most recently, PredMed was developed to extract medication names and related terms from office visit notes.[55] The comparison of PredMed for extracting medication names to earlier versions of MedEx and MedXN on a test set of 50 visit

encounter notes showed F-scores of 80.0% for PredMed, 74.8% for MedEx, and 83.9% for MedXN. Since MedEx-UIMA and MedXN are available as open source systems, we used these systems as baselines for comparison in our study.

In another related work, Kim et al. developed a method for retrieval of biomedical terms in tele-health call notes.[56] Their team identified two types of noise in these records, explicit – including “spelling errors, unfinished sentences, omission of sentence marks, etc.” – and implicit – “non-patient information and a patient’s untrustworthy information” – and sought to remove that noise as part of their method. Utilizing a bootstrapping-based pattern learning process to detect variations related to the explicit noise, and dependency path-based filters to remove the implicit noise, their system achieved an F-score of 77.33% for detecting biomedical terms on evaluation data from 300 patients. This tool and its corresponding code base is not publicly available for comparison for this study. Furthermore, recently, there has been additional work on the analysis of medical conversation.[57–61] However, the corresponding models and methodologies for these studies are not publicly available for evaluation and comparison in this project.

## **Materials and Methods**

Our NLP pipeline was developed and validated to extract medication mentions in clinic visit transcripts. We define medication mentions as any place in the text that a term refers to a medication by a specific or general name or common lay term. Our pipeline takes advantage of Apache clinical Text Analysis and Knowledge Extraction System (cTAKES)[62] to generate a primary candidate list of medication mentions. Subsequently, our approach filters out false positive medical mentions in this list and adds the medication mentions that cTAKES misses in visit transcripts.

### **Visit transcripts dataset**

Transcripts of 85 patient visits with a primary care physician were used as our dataset in this study. These visits were audio recorded and transcribed by a HIPAA compliant commercial medical transcription service. These recordings, which came from eight clinicians, were 31 minutes long on average, ranging from 5.5 to 70.5 minutes. This study and the use of human subject data in this project were approved by the committee for the Protection of

Human Subjects at Dartmouth College (CPHS STUDY#30126) with informed consent. Table 1 shows the demographics of the participants who had their clinical visit recordings used in our study.

Ten transcripts were randomly selected from this dataset as a development set. Another ten of the visit transcripts were randomly selected as a validation set for our model. The remaining 65 transcripts were reserved as a held-out test set for evaluation.

**Table 1.** Participant demographics for transcribed visit recordings (SD: standard deviation).

	<b>Development Dataset</b>	<b>Validation Dataset</b>	<b>Evaluation Dataset</b>	<b>Total</b>
Number of Recordings in Dataset	10	10	65	85
Participants with Demographic Data*	9*	10	54*	73*
Gender				
Female	4	6	34	44
Male	5	4	20	29
Mean age (SD) [Range]	50.00 (18.57) [23-87]	58.60 (18.95) [20-77]	54.65 (15.61) [25-92]	54.62 (16.35) [20-92]
Race				
White	9	10	54	73
Ethnicity				
Not Hispanic or Latino	9	10	52	71
Declines to list	-	-	2	2
Language Spoken				
English	9	10	54	73
Recording length (SD) [Range]	36.46 (17.37) [17.55-70.39]	37.07 (10.00) [20.16-49.41]	28.36 (11.95) [5.42-55.33]	30.55 (12.85) [5.42-70.39]
Visit type				
Annual physical established patient	3	6	8	17
Established patient follow-up	2	3	29	34
Same day add-on	2	1	11	14
New patient workup	2	-	1	3
History and physical	-	-	2	2
Other#	-	-	3	3

\*Demographic data was not captured for 12 of the 85 transcripts.

# “Other” includes “Res-visit 20” and diabetic follow-up.

### **Annotation for medication mentions**

All the transcripts were independently annotated for medication mentions by two second-year medical students using the Extensible Human Oracle Suite of Tools (eHOST) software. [63] The two annotators initially worked through blocks of 5 or 10 transcripts, meeting after annotating each block to track inter-annotator agreement (IAA) on the identified medication mentions, discuss disagreements, and improve their accuracy in this annotation task,

which led to steadily higher IAA over time. Our IAA calculation considers overlapping annotations as a match, allowing a flexible annotation arrangement for compound medication names. Once the annotators reached over 80% IAA, we considered them trained in this annotation task. Subsequently, they annotated the entire set of transcripts. Inter-annotator agreement for medication mentions between our annotators for the 65 transcripts in the evaluation dataset was 84.6%. In that dataset, Annotator 1 annotated 1,076 instances of medication mentions, and Annotator 2 annotated 1,048 instances of medication mentions.

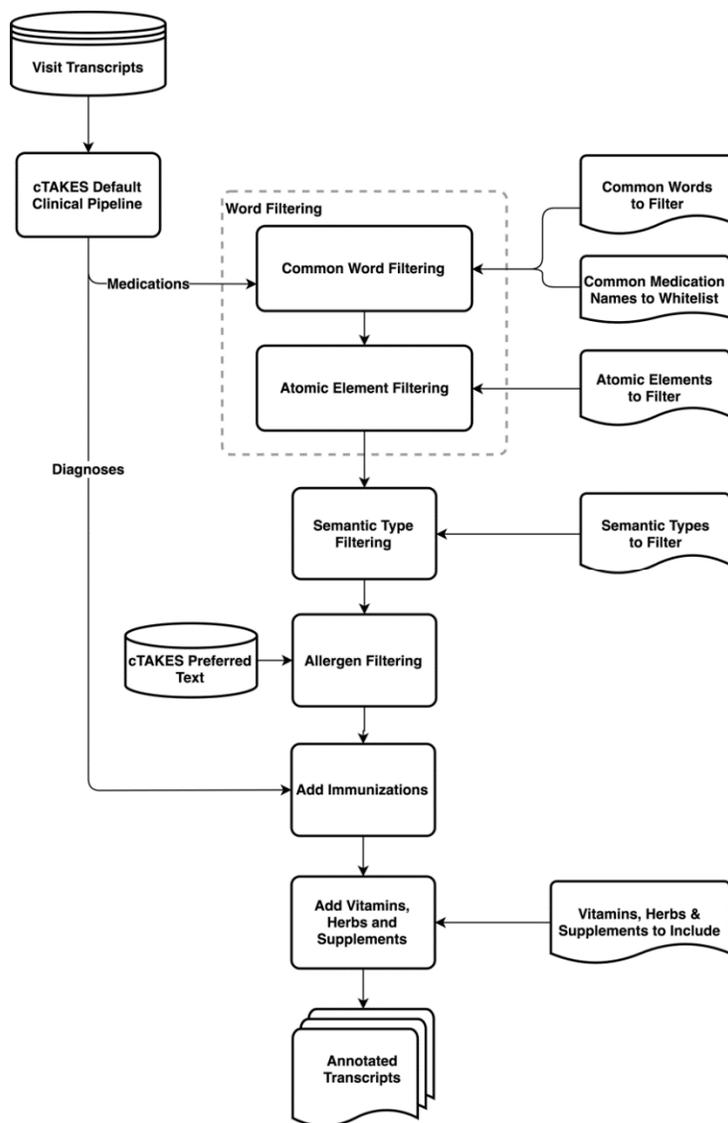
### **cTAKES baseline for annotating medications in transcripts**

Our baseline approach was to utilize Apache cTAKES[62] to identify the medication mentions in the transcripts. cTAKES is an open-source widely-used NLP system for biomedical text processing. As one of its NLP capabilities, cTAKES is able to annotate and extract medical information from the free text of clinical reports. In its default configuration (i.e., Default Clinical Pipeline), cTAKES identifies parts of speech (POS) in input sentences, and annotates concepts found in the noun phrases that match with concepts from the Unified Medical Language System (UMLS) Metathesaurus.[64] We utilized the Default Clinical Pipeline of cTAKES (version 4.0.0) and its UMLS fast dictionary lookup functionality. The only modification to the default cTAKES configuration was to utilize its PrecisionTermConsumer function, which refines annotations to the most specific variation (e.g., if it finds the text “colon cancer” in a report, it only annotates “colon cancer” but not “colon” nor “cancer”). Since cTAKES is designed to work with medical record free text, there is an assumption that input text is a clinical note, written by an individual with a medical background. In contrast, the visit transcripts are typically a dyadic conversation between a patient and their physician.

### **Our model for annotating medications in transcripts**

After initial experiments with cTAKES and UMLS as a means to find medications mentioned in transcribed clinic visit conversations, we explored additional methods to filter out common false positives from the output generated by cTAKES. For this purpose, we took an iterative approach, looking at the most common errors in cTAKES outcomes for identification of medication mentions in our development set and developed new rule-based filters to detect and remove those from the cTAKES output. As our accuracy on the development set improved by filtering out many types of false positives (described in detail below), we ran our model against our validation set, finding that immunizations along with herbs and supplements persisted as typical errors. cTAKES had difficulty

differentiating immunizations from diagnoses (e.g., chickenpox vaccine vs. chickenpox). Also, cTAKES did not annotate some commonly used herbs and supplements. In the next sections, we describe how our approach adds annotations for immunizations, herbs, and supplements, while filtering out false positives for medication mentions. An overview of this approach is shown in Figure 1. We have made our code for this approach publicly available on GitHub.<sup>1</sup>



**Figure 1.** Overview of our approach to annotate medication mentions in clinic visit transcripts.

<sup>1</sup> <https://github.com/BMIRDS/HealthTranscriptAnnotator>

## Common Word Filtering

Since many of the words appearing as false positives in the cTAKES output for medication annotations are common conversational words that have second meanings as medication names or acronyms (for example, “today” is also ToDAY, a name for an antibiotic primarily in veterinary use that appears in UMLS), we decided to utilize a large dictionary of common words to filter out these occurrences. We chose to use a dictionary of the 10,000 most common English words from Google's Trillion Word Corpus<sup>2</sup>. [65] If any of those 10,000 words were annotated by cTAKES as a medication, our model removes that annotation, with a small subset of exceptions. From the 10,000 common words list, there were 24 words that are considered as exceptions and are allowed to remain annotated as medications. These words fit into three categories: 1) names of common medications (e.g., “Ambien”, “Insulin”, etc., which accounted for 17 of the 24); 2) generic terms (e.g., “herb”, “supplement”, and “vitamin”, along with their plurals); and 3) the word “flu”, which can refer to either a diagnosis or an immunization.

## Chemical Element Filtering

cTAKES also annotates all chemical elements as medication mentions; “gold”, for example, can also be taken as a medication. In our approach, we systematically remove annotations for those chemical elements that are not typically taken as a medication or as a supplement.

## UMLS Semantic Type Filtering

In our error analysis for cTAKES outputs, we also examined UMLS semantic types for the terms that cTAKES annotated as medication mentions. The six types shown in Table 2 generally produced false positives and few to no true positives. Our approach removes these semantic types as medication annotations from the cTAKES output where they occur.

**Table 2.** UMLS semantic types in cTAKES annotations that are filtered out in our approach.

<b>TUI</b>	<b>Semantic Type</b>
T114	Nucleic Acid, Nucleoside, or Nucleotide
T122	Biomedical or Dental Material
T123	Biologically Active Substance
T125	Hormone
T130	Indicator, Reagent, or Diagnostic Aid
T197	Inorganic Chemical

---

<sup>2</sup> <https://github.com/first20hours/google-10000-english>

## Allergen Filtering

cTAKES annotates a number of food and food ingredient related terms (for example, “coconut”) as medication mentions, denoting them as an allergenic. We identify those annotations that have the word “allergenic” included in their preferred cTAKES text metadata, and we remove those annotations from the output of cTAKES in our model’s output.

## Immunization Additions

A small number of medication-related UMLS terms are considered as both diagnoses and immunizations/vaccinations (for example, “flu” and “pertussis”). As a result, cTAKES annotation outputs were inconsistent about annotating these terms as immunizations/vaccinations or diagnoses. To improve the annotation of immunizations as medications, we also investigated the cTAKES diagnosis annotations. Since cTAKES segments the input text into sentences, we searched for the words “vaccine”, “shot”, “booster” and “pill” in the same sentence as a diagnosis annotation, and if both co-occurred, we annotated the diagnosis text as a medication.

## Vitamin, Herb, and Supplement Additions

cTAKES also produces inconsistent results for annotating herbs and supplements. Our approach adds an additional dictionary of common herbs and supplements from MedlinePlus<sup>3</sup> to capture these.[66]

## Evaluation

We applied our model on the evaluation dataset containing 65 transcripts to annotate medication mentions, in addition to capturing the original medication mention annotation output from cTAKES 4.0.0’s default clinical pipeline. We also applied publicly available MedEx-UIMA 1.3.7 and MedXN 1.0.1 software on the evaluation dataset to compare our results with their medication name annotations as the baselines.

For evaluation, we created a set of gold standard medication mentions in our evaluation dataset based on the work of our expert annotators. All medication mentions in our evaluation set that were agreed upon by the two expert annotators were kept in this gold standard set. A physician, trained in the method used by the annotators, served as an adjudicator to resolve disagreements between our annotators. A disagreement in the annotations would occur when one annotator had annotated a medication mention while the other had not. Disagreements were resolved by

---

<sup>3</sup> [https://medlineplus.gov/druginfo/herb\\_All.html](https://medlineplus.gov/druginfo/herb_All.html)

the adjudicating physician either choosing to keep the annotation from a single annotator in the gold standard set, or choosing to reject it. The adjudicating physician also reviewed disagreements between the output from our model and the set of annotations from human experts. Again, the adjudicator could either choose to keep the annotation from either source in the gold standard set or choose to reject it. As a result, a small number of medication mentions that were missed by both annotators were thus added to our gold standard set. The resulting gold standard evaluation dataset contained 1,121 medication mentions.

## Results

We calculated the standard evaluation metrics of precision, recall, and F-score for our proposed approach and the baseline methods using the medication mention gold standards in our evaluation set. We compared the results from cTAKES, MedEx-UIMA, MedXN, and our proposed model for identification of the gold standard medication mentions for the 65 transcripts in the evaluation set. Table 3 shows this comparison.

**Table 3.** The comparison of our proposed approach to existing baseline models for identification of gold standard medication mentions in our evaluation set.

Model	# True Positives	# False Positives	# False Negatives	Precision	Recall	F-Score
cTAKES	1,119	2,814	163	28.5%	<b>87.3%</b>	42.9%
MedEx-UIMA	830	1,215	292	40.6%	74.0%	52.4%
MedXN	832	318	432	72.3%	65.8%	68.9%
Our Approach	1,062	168	206	<b>86.3%</b>	83.8%	<b>85.0%</b>

## DISCUSSION

Our results indicate that the proposed approach significantly reduced the number of false positives, with a relatively small drop in the number of true positives and false negatives, in comparison to the best of three baseline models. As highlighted in Table 3, our proposed model has the best overall performance in comparison to the other baseline methods, with all of its evaluation metrics falling in the range of 83% to 87%. Overarching the finer aspects of our work is the observation that extracting medical terms from conversational dialogue between patients and their primary care physician has distinct challenges in comparison to extracting terms from typical clinical, note-like reports. To the best of our knowledge, the proposed work in this paper is the first attempt to extract medical terminology from conversations between a patient and their physician. Prior work for finding medication mentions has focused on written clinical reports.<sup>47-51</sup>

Our error analysis suggests that baseline approaches, which rely on dictionaries, struggle with patient-clinician conversational text because of language like filler words (for example, “aha” and “hmm”) matching with abbreviations for medications, and the fact that common conversational words are often used as medication names. We also observed, among the filters that we applied to the original cTAKES outputs, that filtering out “hormone” semantic type had the most impact on the improvement of the results. A slight but consistent majority of false positives in our dataset were from the discussion of lab test results, which will be a focus of our future work to improve the current results.

One advantage of our approach is that each portion of our pipeline was designed to generalize addressing issues seen during development, so our approach was able to recognize terms outside the development/validation datasets. Other rule-based and dictionary-based systems have often relied on whitelisting/blacklisting terms from their development datasets, which limits how they generalize outside their development data. For example, our use of the 10,000 most common English words from Google's Trillion Word Corpus allows us to recognize and filter many common words. Our solutions to chemical elements, UMLS semantic types, allergens, immunizations, and vitamins/herbs/supplements were also all designed to potentially recognize terms outside what appeared in our development dataset.

Our evaluation has some limitations. Foremost, our evaluation dataset is relatively small and is from a single medical institution. We plan to extend our evaluation dataset in future work to test the generalizability of the proposed approach. In addition, because our gold standard was created by reaching consensus between two medical annotators and carrying out our approach, it is possible that other baseline methods, such as cTAKES, found a small number of true positives that were not accounted for by any of the annotators or our proposed method. That said, the sheer number of false positives generated by cTAKES makes adjudication of its medication mention output impractical. Finally, our approach, which is based on controlled vocabulary and rule-based filtering, does not consider word context and the corresponding contextual semantics in different circumstances. Since one of our goals is using these annotations to index segments of clinic visit conversations for end-users to review post-visit, we plan to conduct future work with end-users to determine how these limitations may impact the usability of the system.

Of note, as we fine-tuned our model on the validation set, we observed that context words in a sentence can be critical in our task, for example, for determining mentions of immunizations/vaccinations. Our result suggests that

although the dictionary- and rule-based methods can achieve a promising result (F-score = 85%) for identification of medication mentions in clinic visit conversations, additional improvements in this domain will be gained through considering contextual semantics and machine learning models, which our team will pursue in future work.

## **CONCLUSION**

In this work, we developed an NLP pipeline for finding medication mentions in primary care visit conversations. The proposed model achieved promising results (Precision = 86.3%, Recall = 83.8%, F-Score = 85.0%) for identification of medication mentions in 65 clinic visit transcripts in our evaluation set. Since this is a first-of-a-kind study with clinic visit transcripts, we compared our approach to three existing systems used for extracting medication mentions from clinical notes. This comparison shows our approach can extract about 27% more medication mentions while eliminating many false positives in comparison to existing baseline systems. Integration of this annotation system with clinical recording applications has the potential to improve patients' understanding and recall of key information from their clinic visits, and, in turn, behavioral and health-related outcomes. We plan to explore this potential in future trials of our system.

## **ACKNOWLEDGEMENTS**

The authors would like to thank Sheryl Piper and Roger Arend for their input on the project and thank Lamar Moss for his feedback on the manuscript.

## **COMPETING INTERESTS**

No conflicts to declare: CHG, WW, PJB, WH, MDD, KLB, JCF, JAS, WMO, JR, MLB, AKD, SH.

GE: Glyn Elwyn has edited and published books that provide royalties on sales by the publishers: the books include Shared Decision Making (Oxford University Press) and Groups (Radcliffe Press). Glyn Elwyn's academic interests are focused on shared decision making and coproduction. He owns copyright in measures of shared decision making and care integration, namely collaboRATE, integRATE (measure of care integration, consideRATE (patient experience of care in serious illness), coopeRATE (measure of goal setting), incorpoRATE (clinician attitude to shared decision making, Observer OPTION-5 and Observer OPTION-12 (observer measures of shared decision making). He has in the past provided consultancy for organizations, including: 1) Emmi Solutions LLC who

developed patient decision support tools; 2) National Quality Forum on the certification of decision support tools; 3) Washington State Health Department on the certification of decision support tools; 4) SciMentum LLC, Amsterdam (workshops for shared decision making). He is the Founder and Director of &think LLC which owns the registered trademark for Option Grids™ patient decision aids; Founder and Director of SHARPNETWORK LLC, a provider of training for shared decision making. He provides advice in the domain of shared decision making and patient decision aids to: 1) Access Community Health Network, Chicago (Adviser to Federally Qualified Medical Centers); 2) EBSCO Health (Consultant); 3) Bind On-Demand Health Insurance (Consultant), 4) PatientWisdom Inc (Adviser); 5) abridge AI Inc (Chief Clinical Research Scientist)

## **FUNDING**

Research reported in this publication was supported by the National Library of Medicine of the National Institutes of Health under award number [R01 LM012815] and the Gordon & Betty Moore Foundation under award number [GBMF-4952]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Gordon and Betty Moore Foundation.

## REFERENCES

- 1 Watson PW, Mckinstry B. A systematic review of interventions to improve recall of medical advice in healthcare consultations: *J R Soc Med* Published Online First: 1 September 2009. doi:10.1258/jrsm.2009.090013
- 2 Kessels RPC. Patients' Memory for Medical Information. *J R Soc Med* 2003;**96**:219–22. doi:10.1177/014107680309600504
- 3 Jansen J, Butow PN, van Weert JCM, *et al.* Does Age Really Matter? Recall of Information Presented to Newly Referred Patients With Cancer. *J Clin Oncol* 2008;**26**:5450–7. doi:10.1200/JCO.2007.15.2322
- 4 Ley P. Memory for medical information. *Br J Soc Clin Psychol* 1979;**18**:245–55. doi:10.1111/j.2044-8260.1979.tb00333.x
- 5 Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract ECP* 1998;**1**:2–4.
- 6 Hibbard JH, Greene J. What The Evidence Shows About Patient Activation: Better Health Outcomes And Care Experiences; Fewer Data On Costs. *Health Aff* 2013;**32**:207–14. doi:10.1377/hlthaff.2012.1061
- 7 Bayliss EA, Ellis JL, Steiner JF. Barriers to Self-Management and Quality-of-Life Outcomes in Seniors With Multimorbidities. *Ann Fam Med* 2007;**5**:395–402. doi:10.1370/afm.722
- 8 Violan C, Foguet-Boreu Q, Flores-Mateo G, *et al.* Prevalence, Determinants and Patterns of Multimorbidity in Primary Care: A Systematic Review of Observational Studies. *PLoS One* 2014;**9**:e102149. doi:10.1371/journal.pone.0102149
- 9 Rocca WA, Boyd CM, Grossardt BR, *et al.* Prevalence of Multimorbidity in a Geographically Defined American Population: Patterns by Age, Sex, and Race/Ethnicity. *Mayo Clin Proc* 2014;**89**:1336–49. doi:10.1016/j.mayocp.2014.07.010
- 10 Bayliss EA, Bayliss MS, Ware JE, *et al.* Predicting declines in physical function in persons with multiple chronic medical conditions: What we can learn from the medical problem list. *Health Qual Life Outcomes* 2004;**2**:47. doi:10.1186/1477-7525-2-47
- 11 Condelius A, Edberg A-K, Jakobsson U, *et al.* Hospital admissions among people 65+ related to multimorbidity, municipal and outpatient care. *Arch Gerontol Geriatr* 2008;**46**:41–55. doi:10.1016/j.archger.2007.02.005
- 12 Bopp KL, Verhaeghen P. Aging and Verbal Memory Span: A Meta-Analysis. *Journals Gerontol Ser B* 2005;**60**:P223–P233. doi:10.1093/geronb/60.5.P223
- 13 Brown SC, Park DC. Roles of Age and Familiarity in Learning Health Information. *Educ Gerontol* 2002;**28**:695–710. doi:10.1080/03601270290099723
- 14 Grady CL, Craik FIM. Changes in memory processing with age. *Curr Opin Neurobiol* 2000;**10**:224–31. doi:10.1016/S0959-4388(00)00073-8
- 15 McCarthy DM, Waite KR, Curtis LM, *et al.* What Did the Doctor Say? Health Literacy and Recall of Medical Instructions. *Med Care* 2012;**50**:277–82. doi:10.1097/MLR.0b013e318241e8e1
- 16 Kutner M, Greenberg E, Jin Y, *et al.* The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy. 2006. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006483>
- 17 Lehnert T, Heider D, Leicht H, *et al.* Review: Health Care Utilization and Costs of Elderly Persons With Multiple Chronic Conditions: *Med Care Res Rev* Published Online First: August 2011. doi:10.1177/1077558711399580
- 18 Boyd CM, Darer J, Boulton C, *et al.* Clinical Practice Guidelines and Quality of Care for Older Patients With Multiple Comorbid Diseases: Implications for Pay for Performance. *JAMA* 2005;**294**:716–24. doi:10.1001/jama.294.6.716
- 19 Hughes LD. Using Clinical Practice Guidelines in Multimorbid Older Adults—A Challenging Clinical Dilemma. *J Am Geriatr Soc* 2012;**60**:2180–2. doi:10.1111/j.1532-5415.2012.04223.x
- 20 Fortin M, Bravo G, Hudon C, *et al.* Relationship Between Multimorbidity and Health-Related Quality of Life of Patients in Primary Care. *Qual Life Res* 2006;**15**:83–91. doi:10.1007/s11136-005-8661-z
- 21 Fortin M, Lapointe L, Hudon C, *et al.* Multimorbidity and quality of life in primary care: a systematic review. *Health Qual Life Outcomes* 2004;**2**:51. doi:10.1186/1477-7525-2-51
- 22 Fortin M, Bravo G, Hudon C, *et al.* Psychological Distress and Multimorbidity in Primary Care. *Ann Fam Med* 2006;**4**:417–22. doi:10.1370/afm.528
- 23 Marengoni A, Winblad B, Karp A, *et al.* Prevalence of Chronic Diseases and Multimorbidity Among the Elderly Population in Sweden. *Am J Public Health* 2008;**98**:1198–200. doi:10.2105/AJPH.2007.121137
- 24 Moffat K, Mercer SW. Challenges of managing people with multimorbidity in today's healthcare systems.

- BMC Fam Pract* 2015;**16**:129. doi:10.1186/s12875-015-0344-4
- 25 McPhail SM. Multimorbidity in chronic disease: impact on health care resources and costs. *Risk Manag. Healthc. Policy*. 2016. doi:10.2147/RMHP.S97248
- 26 Wallace E, Salisbury C, Guthrie B, *et al*. Managing patients with multimorbidity in primary care. *BMJ* 2015;**350**. doi:10.1136/bmj.h176
- 27 Hummel J, Evans P. Providing clinical summaries to patients after each office visit: a technical guide. 2012. <http://hit.qualishealth.org/sites/default/files/hit.qualishealth.org/Providing-Clinical-Summaries-0712.pdf>
- 28 Emani S, Healey M, Ting DY, *et al*. Awareness and Use of the After-Visit Summary Through a Patient Portal: Evaluation of Patient Characteristics and an Application of the Theory of Planned Behavior. *J Med Internet Res* 2016;**18**:e77. doi:10.2196/jmir.5207
- 29 Pavlik V, Brown AE, Nash S, *et al*. Association of Patient Recall, Satisfaction, and Adherence to Content of an Electronic Health Record (EHR)–Generated After Visit Summary: A Randomized Clinical Trial. *J Am Board Fam Med* 2014;**27**:209–18. doi:10.3122/jabfm.2014.02.130137
- 30 Gaston CM, Mitchell G. Information giving and decision-making in patients with advanced cancer: A systematic review. *Soc Sci Med* 2005;**61**:2252–64. doi:10.1016/j.socscimed.2005.04.015
- 31 Weiss L, Gany F, Rosenfeld P, *et al*. Access to Multilingual Medication Instructions at New York City Pharmacies. *J Urban Heal* 2007;**84**:742. doi:10.1007/s11524-007-9221-3
- 32 Tsulukidze M, Durand M-A, Barr PJ, *et al*. Providing recording of clinical consultation to patients – A highly valued but underutilized intervention: A scoping review. *Patient Educ Couns* 2014;**95**:297–304. doi:10.1016/j.pec.2014.02.007
- 33 Good DW, Delaney H, Laird A, *et al*. Consultation audio-recording reduces long-term decision regret after prostate cancer treatment: A non-randomised comparative cohort study. *Surg* 2016;**14**:308–14. doi:10.1016/j.surge.2014.10.006
- 34 Ford S, Fallowfield L, Hall A, *et al*. The influence of audiotapes on patient participation in the cancer consultation. *Eur J Cancer* 1995;**31**:2264–9. doi:10.1016/0959-8049(95)00336-3
- 35 Scott JT, Entwistle VA, Sowden AJ, *et al*. Giving tape recordings or written summaries of consultations to people with cancer: a systematic review. *Heal Expect* 2001;**4**:162–9. doi:10.1046/j.1369-6513.2001.00127.x
- 36 McClement SE, Hack TF. Audio-taping the oncology treatment consultation: a literature review. *Patient Educ Couns* 1999;**36**:229–38. doi:10.1016/S0738-3991(98)00095-0
- 37 Krackow KA, Buyea CM. Use of Audiotapes for Patient Education, Medical Record Documentation, and Informed Consent in Lower Extremity Reconstruction. *Orthopedics* 2001;**24**:683–5. doi:10.3928/0147-7447-20010701-22
- 38 Santo A, Laizner AM, Shohet L. Exploring the value of audiotapes for health literacy: a systematic review. *Patient Educ Couns* 2005;**58**:235–43. doi:10.1016/j.pec.2004.07.001
- 39 Elwyn G, Barr PJ, Grande SW. Patients recording clinical encounters: a path to empowerment? Assessment by mixed methods. *BMJ Open* 2015;**5**:e008566. doi:10.1136/bmjopen-2015-008566
- 40 Tsulukidze M, Grande SW, Thompson R, *et al*. Patients Covertly Recording Clinical Encounters: Threat or Opportunity? A Qualitative Analysis of Online Texts. *PLoS One* 2015;**10**:e0125824. doi:10.1371/journal.pone.0125824
- 41 Barr PJ, Bonasia K, Verma K, *et al*. Audio-/Videorecording Clinic Visits for Patient’s Personal Use in the United States: Cross-Sectional Survey. *J Med Internet Res* 2018;**20**:e11308. doi:10.2196/11308
- 42 Tang PC, Ash JS, Bates DW, *et al*. Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. *J Am Med Informatics Assoc* 2006;**13**:121–6. doi:10.1197/jamia.M2025
- 43 Bayati M, Braverman M, Gillam M, *et al*. Data-Driven Decisions for Reducing Readmissions for Heart Failure: General Methodology and Case Study. *PLoS One* 2014;**9**:e109264. doi:10.1371/journal.pone.0109264
- 44 Hassanpour S, Langlotz CP. Predicting High Imaging Utilization Based on Initial Radiology Reports: A Feasibility Study of Machine Learning. *Acad Radiol* 2016;**23**:84–9. doi:10.1016/j.acra.2015.09.014
- 45 Hassanpour S, Langlotz CP. Information extraction from multi-institutional radiology reports. *Artif Intell Med* 2016;**66**:29–39. doi:10.1016/j.artmed.2015.09.007
- 46 Hassanpour S, Bay G, Langlotz CP. Characterization of Change and Significance for Clinical Findings in Radiology Reports Through Natural Language Processing. *J Digit Imaging* 2017;**30**. doi:10.1007/s10278-016-9931-8
- 47 Huhdanpaa HT, Tan WK, Rundell SD, *et al*. Using Natural Language Processing of Free-Text Radiology Reports to Identify Type 1 Modic Endplate Changes. *J Digit Imaging* 2018;**31**. doi:10.1007/s10278-017-0013-3

- 48 Meng X, Ganoë CH, Sieberg RT, *et al.* Assisting radiologists with reporting urgent findings to referring physicians: A machine learning approach to identify cases for prompt communication. *J Biomed Inform* 2019;**93**. doi:10.1016/j.jbi.2019.103169
- 49 Hassanpour S, Langlotz CP. Unsupervised Topic Modeling in a Large Free Text Radiology Report Repository. *J Digit Imaging* 2016;**29**. doi:10.1007/s10278-015-9823-3
- 50 Barr PJ, Dannenberg MD, Ganoë CH, *et al.* Sharing Annotated Audio Recordings of Clinic Visits With Patients—Development of the Open Recording Automated Logging System (ORALS): Study Protocol. *JMIR Res Protoc* 2017;**6**:e121. doi:10.2196/resprot.7735
- 51 Uzuner Ö, Solti I, Cadag E. Extracting medication information from clinical text. *J Am Med Informatics Assoc* 2010;**17**:514–8. doi:10.1136/jamia.2010.003947
- 52 Patrick J, Li M. High accuracy information extraction of medication information from clinical notes: 2009 i2b2 medication extraction challenge. *J Am Med Informatics Assoc* 2010;**17**:524–7. doi:10.1136/jamia.2010.003939
- 53 Sohn S, Clark C, Halgrim SR, *et al.* MedXN: an open source medication extraction and normalization tool for clinical text. *J Am Med Informatics Assoc* 2014;**21**:858–65. doi:10.1136/amiajnl-2013-002190
- 54 Jiang M, Wu Y, Shah A, *et al.* Extracting and standardizing medication information in clinical text – the MedEx-UIMA system. *AMIA Summits Transl Sci Proc* 2014;**2014**:37–42. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4419757/>
- 55 Wang Y, Steinhubl SR, Defilippi C, *et al.* Prescription Extraction from Clinical Notes: Towards Automating EMR Medication Reconciliation. *AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci* 2015;**2015**:188–93.
- 56 Kim M-Y, Xu Y, Zaiane OR, *et al.* Recognition of Patient-Related Named Entities in Noisy Tele-Health Texts. *ACM Trans Intell Syst Technol* 2015;**6**:59:1–59:23. doi:10.1145/2651444
- 57 Selvaraj SP, Konam S. Medication Regimen Extraction from Medical Conversations. *Stud Comput Intell* 2021;**914**:195–209. doi:10.1007/978-3-030-53352-6\_18
- 58 Patel D, Konam S, Selvaraj SP. Weakly Supervised Medication Regimen Extraction from Medical Conversations. Published Online First: 11 October 2020. <http://arxiv.org/abs/2010.05317> (accessed 24 Mar 2021).
- 59 Mani A, Palaskar S, Konam S. Towards Understanding ASR Error Correction for Medical Conversations. 2020;:7–11. doi:10.18653/v1/2020.nlpmc-1.2
- 60 Du N, Wang M, Tran L, *et al.* Learning to infer entities, properties and their relations from clinical conversations. *EMNLP-IJCNLP 2019 - 2019 Conf Empir Methods Nat Lang Process 9th Int Jt Conf Nat Lang Process Proc Conf* 2020;:4979–90. doi:10.18653/v1/d19-1503
- 61 Enarvi S, Amoia M, Del-Agua Teba M, *et al.* Generating Medical Reports from Patient-Doctor Conversations Using Sequence-to-Sequence Models. 2020;:22–30. doi:10.18653/v1/2020.nlpmc-1.4
- 62 Savova GK, Masanz JJ, Ogren P V, *et al.* Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Informatics Assoc* 2010;**17**:507–13. doi:10.1136/jamia.2009.001560
- 63 South BR, Shen S, Leng J, *et al.* A Prototype Tool Set to Support Machine-assisted Annotation. In: *Proceedings of the 2012 Workshop on Biomedical Natural Language Processing*. Stroudsburg, PA, USA: : Association for Computational Linguistics 2012. 130–9. <http://dl.acm.org/citation.cfm?id=2391123.2391141>
- 64 Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004;**32**:D267–D270. doi:10.1093/nar/gkh061
- 65 Kaufman J. 10,000 most common English words from Google’s Trillion Word Corpus. 2018. <https://github.com/first20hours/google-10000-english>
- 66 Herbs and Supplements: MedlinePlus. [https://medlineplus.gov/druginfo/herb\\_All.html](https://medlineplus.gov/druginfo/herb_All.html)