Robust Methods for Clinical Text Classification and Disease Understanding with NLP Extracted Symptoms from Clinical Notes

Weipeng Zhou

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Meliha Yetisgen, Chair

Steve Mooney

Timothy Miller

Program Authorized to Offer Degree:

Biomedical Informatics and Medical Education

University of Washington

Abstract

Robust Methods for Clinical Text Classification and Disease Understanding with NLP Extracted Symptoms from Clinical Notes

Weipeng Zhou

Chair of the Supervisory Committee:
Meliha Yetisgen
Department of Biomedical Informatics and Medical Education

Electronic Health Records (EHR) contain comprehensive medical and treatment histories of patients and have the potential to be used to provide better healthcare. A significant portion of the EHR is in the form of clinical notes and Natural Language Processing (NLP) methods can help extract hidden information from them. However, applying NLP in healthcare has challenges. Many of the clinical note datasets are scarce and imbalanced, making it difficult to develop generalizable and robust NLP methods. Additionally, effective use of NLP in healthcare requires close collaboration with medical experts to identify and understand meaningful clinical problems. This dissertation addresses these challenges and explores the application of NLP in healthcare. In Chapter 3 and 4, we develop generalizable and robust NLP methods for clinical note classification and female suicide report coding. In Chapter 5 and 6, we apply NLP to extract

symptoms from clinical notes and study risk factors associated with out-of-hospital cardiac arrest (OHCA) and Long COVID.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would first like to thank Meliha Yetisgen, my PhD advisor who kindly accepted me as a member of her lab in the third year of my PhD, and when I had a background irrelevant to NLP. She has been providing much advice to me on becoming a successful PhD. I was also given much computing resources (remote servers, GPUs, HIPAA compliant ChatGPT server) so I could get hands-on experience with large language models when they first came out. Meliha also connected me with various NLP and clinical collaborators, such as Timothy Miller and Kari Stephens. In this process I learned about communicating with different people and they became an indispensable part to the achievements I made during my PhD.

I would also like to thank Steve Mooney, who I have been working as a research assistant with since the second year of my PhD. Together, we developed software for gathering neighborhood context measures (e.g., education level) from scratch, maintained it for four years, and helped install and run the software for researcher groups across the United States (with the help of Amy Youngbloom who is also a critical developer of the software).

I would like to thank Timothy Miller, who I interned with in the summer of 2022 and 2023, but actually had weekly meetings over the past two years. With Tim's help, I published my first JAMIA paper, which helped me gather a lot confidence and contributed to my decision of pursuing a PostDoc position after graduation.

Lastly, I would like to thank Kari Stephens, who is not on my dissertation committee, but helped me significantly in writing and implementing a pilot grant for Long COVID studies, and this work eventually became a chapter of this thesis.

Chapter 1. INTRODUCTION

1.1 PROBLEM

Electronic Health Records (EHRs) contain comprehensive patient medical history and treatment details[1]. Machine learning methods play a crucial role by not only predicting health outcomes but providing a deeper understanding of diseases and more tailored treatment plans for patients[2,3]. By Utilizing the EHR data, people can develop high-performance machine learning methods. For example, In Futoma et al.[4], researchers used neural network to predict hospital readmission for discharged patients and in Bajor et al.[5], researchers were able to predicted sequential events of medications based on the patients' previous treatment records. Nevertheless, a lot of EHR data are buried in text, such as clinical notes and radiology reports[6], and uncovering them using natural language processing (NLP) to provide a comprehensive representation for the patient has become a major topic.

To date, numerous studies have demonstrated the success of natural language processing (NLP) in meaningful tasks, such as extracting symptom events from clinical notes and enhancing cohort discovery[7,8]. However, challenges remain in employing these models more broadly and achieving a wider impact[2,9,10]. One major issue is the transferability of the trained clinical NLP models. Healthcare institutions document clinical notes in very different ways, making it difficult to apply a model trained in one institution to another. Moreover, many clinical datasets suffer from the problem of data imbalance [11–13], which hinders the delivery of more robust NLP models. Furthermore, despite the rapid development of state-of-the-art NLP methods, their potential in improving healthcare and providing insights to pressing diseases, like cardiac arrest and Long

COVID, is still under explored. Understanding how NLP can provide clinical insights is of critical importance for the future integration of NLP with healthcare.

This dissertation aims to address of these challenges by 1) studying generalizability of classification models for clinical note classification 2) applying novel methods to address the data imbalance issue in coding female suicide report, and 3) using NLP assisted approaches to investigate two pressing diseases and provide public health insights: out-of-hospital cardiac arrest and Long COVID.

1.2 CONTRIBUTIONS

The primary contributions of this work include:

Transferability of clinical note section classification models:

Section classification refers to the task of assigning section names to clinical note segments[14]. It is known as a note preprocessing step that benefits downstream tasks, such as named entity recognition[9] and cohort discovery[15], for it reducing the note size and removing irrelevant information. Current state-of-the-art section classification often achieved high accuracy in-domain but degenerated significantly when transferred to another health care institution, since the note types, note taking habits and section interpretations could be very different across institutions[16]. This work studied methods for improving the transferability of clinical note section classification models across healthcare institutions. We studied three datasets coming from different sources and of different note types. We attempted to enhanced model transferability by applying continued pretraining, which involves pretraining a model on an unannotated corpus of texts[17], and by providing multiple key insights for people willing to develop transferrable section classification models.

Addressing data imbalance - coding highly rare circumstances in female firearm suicide reports with large language models:

Female firearm suicide reports contain valuable information for understanding the recently fast-increasing female suicides in the United States[18]. Nevertheless, the textual reports are unstructured information and not optimal for analysis. Coding (classifying) and assigning circumstance labels (e.g., sleep problems, bullying) to the suicide reports would enable further research. Existing studies developed high performance models for the prevalent circumstances in the reports (e.g., mental problems) but failed for the rarer ones (e.g., sexual violence), which are also important for understanding female suicides. This work addressed this problem by developing novel NLP methods for coding the rare circumstances, where the traditionally machine learning methods struggled due to a sheer data imbalance issue (5% of positives and 95% of negatives). Large language models have demonstrated remarkable performance in generalizing to unknown domains without any annotations[19]. We compared large language models with traditional machine learning methods and suggested novel paradigm for related tasks[20].

Exploring symptom trajectories prior to Out-of-hospital cardiac arrest:

Out-of-hospital cardiac arrest (OHCA), known for its sudden onset and high mortality rate[21], accounts for more than 15% of the global deaths. Current studies of out-of-hospital cardiac arrest rely on observations from bystanders and survivor narratives[22–25], and could lead to observation or recall bias. Besides, such studies limited the data collection period to be hours or days prior to the onset of OHCA, precluding the symptom trajectory analysis that could be extended to weeks or months prior to the onset. Moreover, it is found that a large number of clinically relevant symptoms are buried in clinical notes and not recorded as ICD (International Classification of Disease) codes, making the traditional EHR analytic approaches under representative[26,27].

In addressing this, we proposed a new paradigm for studying the symptoms that could be correlated with OHCA, focusing on symptoms extracted from clinical notes. We collected a comprehensive EHR dataset for OHCA patients and analyzed their symptom trajectories in 12 months prior to OHCA using symptoms extracted from clinical narratives. The trajectories were compared with those of controls, and provided insights into understanding factors leading to OHCA onset. We also developed a machine learning model for predicting the likelihood of patients having OHCA in three months.

Healthcare disparities among Long COVID patient populations

Long COVID refers to the condition where a patient continues to experience symptoms for more than three months after recovering from the acute phase of COVID-19, and the symptoms are not attributable to other causes[28]. Despite numerous studies on Long COVID[29,30], its definition remains vague and lacks a clear definition. One contributing factor to the current status is the evolving understanding of Long COVID, which causes the definition to change over time[29,30]. This evolving definition can lead to healthcare disparities for patients receiving care from different sources because the varying presentations of Long COVID can complicate case identification and consistency in referrals in addition to pre-existing disparities in access to care (e.g., those driven by social determinant of health).

At the University of Washington Medical Center, Long COVID patients may seek care at either primary care clinics or a specialized Long COVID clinic (Post-COVID-19 Rehabilitation and Recovery Clinic). Due to the limited capacity of the specialized clinics, many patients are managed in primary care settings. The quality and scope of care between these two types of clinics can differ significantly, influenced by factors such as the severity of the patient's symptoms and the expertise of the physician.

Without thorough investigation, it is challenging to determine whether healthcare resources are being appropriately allocated to those most in need. Specialized Long COVID clinic offer the latest advances in Long COVID care, which may not be equitably accessible for all patients in need of this stepped-up care. A significant obstacle to such analyses is the under-documentation of symptoms in the its structured format, particularly for COVID-related symptoms[7].

In this study, we compare patients receiving Long COVID care in primary care clinics to those receiving care in a specialized Long COVID clinic. The analysis is done across demographics and Long COVID related symptoms (i.e., ICD codes, ICD code groups [grouped into comorbidities, Long COVID symptoms, and psychiatric symptom], and symptoms extracted from clinical notes).

Chapter 2. BACKGROUND

2.1 CLINICAL NOTE SECTION CLASSIFICATION

Clinical note section classification refers to assigning a section name to the segments of a clinical note. It provides high level information for understanding the structure of a note and is found to be beneficial for a number of clinical NLP tasks including named entity recognition[9], abbreviation resolution[31], cohort retrieval[8] and temporal relation extraction[32] .

Previous research on section classification[16,33] has demonstrated that this task can be successfully accomplished for a specific dataset. However, performance significantly decreases when applying a trained system to a different dataset. In the context of clinical note section classification, both statistical methods and advanced pretrained transformers like BERT[34] have shown impressive results for modeling within a single institution[16,33].

In Denny et al., researchers developed the "SecTag" algorithm for identifying sections in clinical notes[35]. The algorithm employs a combination of natural language processing techniques (e.g., word variant recognition, spelling correction) and naive Bayesian scoring methods. The algorithm was applied on thousands of clinical notes randomly sampled from the Vanderbilt University Medical Center and achieved an F1 score of over 0.9. In Li et al., researchers developed a supervised Hidden Markov model for performing section classification on various note types, including primary provider notes, consultation notes, follow-up, and clinical notes with 15 predefined section types[36]. Different from previous studies which classified each section independently, they took into account the sequential aspect of the section arrangement in a note and framed it as a sequence labeling task. In comparison to the baseline that did not use sequential labeling and had an accuracy of 0.7, the proposed method achieved an accuracy of 0.9. In a study focused on classifying SOAP ("Subjective", "Objective", "Assessment" and "Plan") sections for

emergency department reports[37], researchers developed a Support Vector Machine classifier incorporating various features, such as lexical, syntactic, semantic, contextual, and heuristic attributes. They achieved an F1 score of 0.8532. In another study by Rosenthal et al.[33], BERT achieved F1 scores of 0.99 and 0.9 for two datasets with fine-grained section names. In the work of Tepper et al.[16], researchers examined note segmentation and section classification, with a focus on fine-grained section names. They utilized Maximum Entropy Classifiers with detailed features, including capital letters, numbers, blank lines, and previous section names, achieving an F1 score exceeding 0.9 for two discharge summary datasets and one radiology report dataset. However, when transferring these models from one dataset to another, the F1 score decreased to 0.6.

In comparison to our study, the existing literature focused on developing models using human-created annotated note section data but ignored the vast number of unannotated notes that could be valuable for enhancing model performance. They also mostly performed in-domain experiments, modeling and evaluating on the same note types or notes from the same healthcare institution. Besides, when studying cross-domain model transferability, existing literature tended to focus only on unsupervised adaption, where they transferred the trained model directly to the target domain, and ignored supervised domain adaptation, a scenario where the target domain has some annotated data available.

2.2 RARE CIRCUMSTANCES PRECEDING FEMALE FIREARM SUICIDES

Suicide is a significant cause of mortality in the United States. Several risk factors are associated with suicide, including physical and mental health disorders, substance use disorders, previous exposure to violence, and the presence of firearms in households[18,38–40]. Although firearm suicide historically affects males more, there has been an increase of 20% in female firearm suicide rates between 2010 and 2020, surpassing the rate increase observed among males by 8 percentage

points[41]. Besides, the contributing circumstances to female firearm suicide may differ. Only limited studies are done on identifying the risk factors specific to female firearm suicide, partly because circumstantial data is solely available in narrative form within the National Violent Death Reporting System (NVDRS), which serves as the primary source of detailed suicide information in the United States.

In Crosby et al., researchers studied suicide circumstances among females in 16 states in the United States from 2005 to 2016 using structured data provided by NVDRS[42]. The most common preceding factors were mental health problems and prior mental health treatment. They examined suicide circumstances based on year, age group (10–17, 18–39, 40–64, 65+), and race/ethnicity. The top 10 most commonly reported circumstances for female suicides were identified, including current mental health problems, treatment history, depressed mood, alcohol and substance abuse, and intimate partner issues. These circumstances varied by age group and race/ethnicity. Using NVDRS structured data, another study also examined suicide rates among older women in the United States, focusing on the differences in sociodemographic factors and risk factors among those who died by drug overdose, firearms, hanging/suffocation, or other means[43]. The findings revealed that older women who used drug overdose as a means of suicide were typically younger, had a history of previous suicide attempts, mental disorders, and substance abuse. In contrast, they had lower rates of relationship problems and other crises. The study suggests that suicide prevention efforts should include restricting access to prescription medications and firearms for those at risk and providing better mental health and substance abuse treatment, as well as support for managing chronic illnesses. In a prior study[38], it was observed that traditional natural language processing algorithms demonstrated effectiveness in detecting common circumstances preceding female firearm suicide by analyzing the CME (coroners or medical examiners) and LE (law

enforcement) narrative reports from the NVDRS. Nevertheless, it should be noted that the training of a conventional NLP model demands a sizable dataset, which limited the model performance for the rarer preceding circumstances.

In comparison to existing literatures, we focused exclusively on female firearm suicides due to their increasing threat to population health and yet little is understood about them. Instead of relying on the circumstances coded by NVDRS, we developed methods to identify a broader set of circumstances from suicide reports using natural language processing. Additionally, we prioritized the exploration of rarer circumstances (<5% occurrences in the dataset) in contrast to the more common ones and developed effective approaches for their extraction.

2.3 OUT-OF-HOSPITAL CARDIAC ARREST

Out-of-hospital cardiac arrest (OHCA) is a significant public health challenge, and is responsible for around 15-20% of deaths worldwide[21]. Despite advancements in healthcare, the survival rate for OHCA remains as low as 10%[44,45]. To address this issue, the development of effective strategies for early recognition and prevention is essential[46].

Previous studies[23–25,47] have shown that nearly half of OHCA patients experience symptoms in the hours or days leading up to the event, but only a small number seek emergency medical help when these warning signs occur[23]. While this information is valuable for individuals at immediate risk, we lack insights into how symptoms develop over a longer period, which could provide a wider window for risk assessment and intervention. Additionally, the existing literature relies on reports from bystanders or emergency medical services (EMS), which may be influenced by recall bias. Identifying symptoms reported by patients themselves would offer valuable information for both individuals and healthcare providers. Studying symptom trends before OHCA is also

challenging due to the low overall risk in the general population, making traditional prospective observational studies underpowered for this task.

In a recent study involving cardiac arrest patients in the United States, it was found that OHCA patients usually experienced symptoms like dyspnea, chest pain, diaphoresis and seizure-like activity[24]. Similar findings were seen in studies from Japan[47] and Sweden[22], where people with imminent OHCA frequently reported symptoms like difficulty breathing and chest pain. These studies help us understand how to develop better paradigms for preventing OHCA. However, we still need to learn more about how symptoms change over time to better screen and reduce the risk of OHCA.

The existing literatures usually lack adequate longitudinal analysis on symptom progression before OHCA. Moreover, there is a need for a more representative study group for OHCA and more data analytics (using the rich symptom information in clinical narratives) to identify risk factors associated with OHCA. To address these issues, our study linked a population-based OHCA registry maintained by EMS with electronic health records from the University of Washington healthcare system. The linked dataset covers approximately 1.5 million patients. We employed a case-control study design involving 2366 OHCA cases and 23660 controls. We used this combined dataset for analysis and reported the 1-year symptom trajectory for OHCA patients.

2.4 LONG COVID

Despite many published studies detailing negative sequelae from COVID-19[29,30], the post-acute sequelae of COVID-19, also known as Long COVID, lacks a clear definition. Usually, Long COVID[28] refers to the condition where a patient still has persisting symptoms after recovering from COVID-19 for longer than three months and the symptoms are not explainable by other causes. Patients with Long COVID experience symptoms similar to COVID-19, and additional

symptoms such as brain fog, daytime fatigue and tiring easily. Complete capture of symptoms for Long COVID is crucial for understanding the impacts of COVID-19 long-term and delivering effective treatment.

Key symptoms and functional impacts of Long COVID, such as brain fog, difficulty working and fatigue, are often tracked only in clinical notes due to documentation habits of providers and efficiency demands, as well as a lack of easily accessible available codes. Natural language processing methods have been found to have the potential for extracting them[48]. For example, in the task of extracting COVID-19 symptoms from clinical notes, NLP methods were able to capture 64.7% more symptoms than ICD-10 codes or regular expressions[27].

Various studies tried to understand Long COVID from clinical notes and identifying the key characteristics of Long COVID patients. In Wang et al., researchers aimed to create a comprehensive lexicon of symptoms related to Long COVID[49]. They identified 355 symptoms (with 1,520 UMLS concepts and 16,466 synonyms) from clinical notes of 26,117 COVID-19 patients. The notes were collected during the patients' post-acute infection period, and the researchers used a rule-based natural language processing system to extract symptoms. The lexicon was refined through manual review and consolidation of concepts. The most prevalent symptoms included pain, anxiety, depression, and fatigue. In Pfaff et al., researchers aimed to identify potential Long COVID patients using machine learning models developed from electronic health records[50]. The base population consisted of over 1.7 million non-deceased adult patients who had a COVID-19 diagnosis or a positive SARS-CoV-2 test at least 90 days prior. The models were trained using structured features, including ICD codes and demographic records. The model achieved high accuracy in identifying potential Long COVID patients, and the important features were found to be healthcare utilization, patient age, dyspnea, and other diagnosis and medication

information. In Thaweethai et al., researchers developed a symptom definition for Long COVID based on self-reported symptoms obtained from questionnaires[51]. It involved 9,764 participants, with 89% having a history of COVID infection. The analysis identified 37 symptoms that were more common in COVID individuals 6 months or more after infection compared to uninfected individuals.

This study advanced NLP methods in Long COVID studies further by employing an advanced transformer-based symptom extractor developed previously to accurately extract Long COVID symptoms from clinical notes[27]. This study also improved on past research by using a provider identified Long COVID cohort derived from the University of Washington Medicine healthcare system, specifically the specialized Long COVID clinic, Post-COVID-19 Rehabilitation and Recovery Clinic (Long COVID Clinic).

Chapter 3. TRANSFERABILITY OF CLINICAL NOTE SECTION CLASSIFICATION MODELS

3.1 OVERVIEW

In this chapter, we investigated methods for improving the transferability of clinical note section classification models across domains. To fill the gap in creating section classifiers that work well in various situations, we developed domain adaptation techniques within the context of SOAP section classification. In clinical settings, SOAP-style notes are commonly used for patient care documentation[37,52]. The SOAP ("Subjective", "Objective", "Assessment", and "Plan") note is note structuring method for effectively documenting and communicating patient information. The "Subjective" section, such as "Social history," documents the patient's symptoms, concerns, and medical history, while the "Objective" section, such as "Vital signs," documents observable data including physical examinations and diagnostic tests. The "Assessment" section, such as "Discharge diagnosis," documents the physician's analysis and diagnosis. The "Plan" section, such as "Follow-up," documents the potential plans for the patient, including interventions, medications, and follow-up steps. Classifying sections into SOAP categories can help understand information sources in notes. For instance, data related to social determinants of health (SDOH) is often found in the social history section, a "Subjective" SOAP category. Furthermore, advanced NLP models (pretrained transformers) have word processing limits[34], so identifying sections and shorten the inputs to the them can enhance their applicability.

Different types of clinical notes have different section types (e.g., "Discharge instruction" for discharge summaries and "Overnight progress" for progress notes). By simplifying the classification task to SOAP, we can conduct cross-domain experiments and study performance variations when the variable, section type differences between datasets, is removed.

This work used domain adaptation methods to improve the transferability of clinical note section classification models. We employed domain-adaptive pretraining (DAPT) and task-adapted pretraining (TAPT) methods[53]. These methods applied masked language modeling pretraining on target domain data before fine-tuning the model with labeled source domain data. We also explored the impact of including a small amount of labeled data from the target domain in model fine-tuning to assess the relationship between unsupervised and supervised domain adaptation techniques[17].

3.2 DATASET

We utilized three datasets with different types of medical notes from different healthcare systems. The first dataset is referred to as discharge, and it includes discharge summaries from the i2b2 2010 challenge provided by Partners Healthcare and Beth Israel Deaconess Medical Center[16]. The second dataset is referred to as thyme, and it consists of colorectal clinical notes from the THYME (Temporal History of Your Medical Events) corpus from Mayo Clinic[54]. The third dataset is referred to as progress, and it consists of progress notes from MIMIC-III written by healthcare practitioners from the intensive care units[9,11,13].

For each of these datasets, we generated classification instances by selecting sections from the notes. Although all three datasets had section labels, they were not consistent across datasets. Notably, the progress dataset already had section labels mapped to SOAP categories. To ensure consistency for cross-domain experiments and following the SOAP definition guideline[37], an expert physician informaticist mapped the section labels of the thyme and discharge datasets to SOAP labels[55,56]. Sections that did not fit these SOAP categories, like “Comments” and “Administrative,” were labeled as “Others,” thus creating a 5-way classification system.

Table 1 provides information about each dataset's size, average word count, label distribution, and the train/test split ratio. During the SOAP mapping process, we observed that the "Assessment and Plan" section in the progress dataset encompassed both "Assessment" and "Plan" content. Consequently, we mapped these sections to the "Assessment" category, resulting in a "Not Applicable (N/A)" in Table 1. When dividing the dataset into training and test sets, we used a random 0.8/0.2 split for the discharge dataset and followed the original train/test splits for thyme and progress.

Table 1 also highlights a potential challenge when transferring SOAP classifiers between domains, as the distribution of SOAP categories varies significantly. While "Subjective" and "Objective" are consistently the most prevalent categories, "Objective" has the largest count in discharge and progress, whereas "Subjective" is the largest in thyme.

Table 1. Size, average section word count (with standard deviation), and label distribution of the discharge, thyme and progress dataset.

dataset	total section counts	average word count	"Subjective" section Count	"Objective" section count	"Assessment" section count	"Plan" section count	"Others" section count	train/ test split
discharge	1372	61±112	376 (27.4%)	628 (45.8%)	243 (17.7%)	103 (7.5%)	22 (1.6%)	0.8/0.2
thyme	4223	74±121	1878 (44.5%)	1329 (31.5%)	676 (16.0%)	100 (2.4%)	240 (5.7%)	0.73/0.27
progress	13367	46±97	4521 (33.8%)	7039 (52.7%)	787* (5.9%))	N/A	1020 (7.6%)	0.89/0.11

*Assessment and Plan combined

3.3 IN-DOMAIN SECTION CLASSIFICATION

We employed the pretrained transformer framework for creating the section classification model. BioBERT[57] was fine-tuning on the training set of thyme, discharge, and progress datasets. We chose BioBERT because it outperformed BERT in various biomedical NLP tasks[57]. Since other domain-specific BERT variants like BioClinicalBERT[58] were mostly pretrained on MIMIC-III, which contains our progress dataset, we tried to avoid them in the initial fine-tuning to prevent data leakage. Using BioBERT as the baseline model also enables the convenience use of

BioClinicalBERT, a readily available domain-adapted pretraining (DAPT) version of BioBERT (see future sections for details).

We assessed the in-domain classification performance on the three datasets, setting the performance upper bounds for the cross-domain experiments. We fine-tuned the BioBERT model on the training set of the datasets and evaluated it on the test set of the same dataset. For these in-domain experiments, both the source and target domains were the same, so we denoted these experiments as FT_{target} (fine-tuned directly on the target domain data). Table 2, 3 and 4 summarizes the experiment names and settings for the in-domain and the following cross-domain experiments.

Table 2. Description of experiment configurations with discharge being the target domain.

Method	experiment	source domain	target domain	number of target domain labeled samples added to fine-tuning	DAPT corpus	TAPT corpus
In-domain section classification	FT_{target}	discharge	discharge	all	MIMIC-III	discharge training set
Cross-domain section classification	FT_{source}	thyme or progress		0		
Cross-domain section classification with continued pretraining	DAPT + FT_{source}					
	DAPT + TAPT + FT_{source}					
Cross-domain section classification with continued pretraining and target domain labeled data	$FT_{source + target}$			10,20,30,40,50		
	DAPT + $FT_{source + target}$					
	DAPT + TAPT + $FT_{source + target}$					

Table 3. Summarization of experiment configurations with thyme being the target domain.

Method	experiment	source domain	target domain	number of target domain labeled samples added to fine-tuning	DAPT corpus	TAPT corpus
In-domain section classification	FT_{target}	thyme	thyme	all	unlabeled notes in THYME corpus	thyme training set
Cross-domain section classification	FT_{source}	discharge or progress		0		
Cross-domain section classification with continued pretraining	DAPT + FT_{source}					
	DAPT + TAPT + FT_{source}					
Cross-domain section classification with continued pretraining and target domain labeled data	$FT_{source + target}$			10,20,30,40,50		
	DAPT + $FT_{source + target}$					
	DAPT + TAPT + $FT_{source + target}$					

Table 4. Description of experiment configurations with progress being the target domain.

Method	experiment	source domain	target domain	number of target domain labeled samples added to fine-tuning	DAPT corpus	TAPT corpus
In-domain section classification	FT_{target}	progress	progress	All	MIMIC-III	progress training set
Cross-domain section classification	FT_{source}	discharge or thyme		0		
Cross-domain section classification with continued pretraining	DAPT + FT_{source}					
	DAPT + TAPT + FT_{source}					
Cross-domain section classification with continued pretraining and target domain labeled data	$FT_{source + target}$			10,20,30,40,50		
	DAPT + $FT_{source + target}$					
	DAPT + TAPT + $FT_{source + target}$					

During BERT fine-tuning, we utilized a learning rate of $1e^{-5}$, an epoch size of 40, and a batch size of 10. We saved the best model, determined by the highest F1 score on the validation set, for testing. A single run with the best model is reported and we did not do averaging across runs. This has the tradeoff of showing realistic amounts of noise that could be attributed to small numbers of instances, while not giving a stable estimate of the expected change in performance. We discussed more about the impact of this tradeoff in the discussion section. Because of imbalanced SOAP categories across datasets, we used the micro-F1 score for the evaluation

metric. We used an AdamW optimizer implemented by HuggingFace for fine-tuning[59]. All experiments were conducted on a 24GB NVIDIA TITAN RTX GPU with FP16 precision.

3.4 CROSS-DOMAIN SECTION CLASSIFICATION

We evaluated the performance of cross-domain section classification on three datasets. We fine-tuned a model on the training set of one dataset and tested it on the test sets of the other two datasets. For example, if we trained a model on the training set of discharge, we tested it on the test set of thyme and progress. These experiments are referred to as FT_{source}, where the models were fine-tuned on the source domain but tested on the target domain. We maintained the same model hyperparameter settings as in-domain experiments to simulate scenarios where limited resources are available in the target domain.

3.5 CROSS-DOMAIN SECTION CLASSIFICATION WITH CONTINUED PRETRAINING

Recent research suggests that continued pretraining of pretrained language models in the target domain enhances adaptability[53]. There are two types of continued pretraining techniques. Domain-adaptive pretraining (DAPT) involves further training a pretrained model on a large collection of unlabeled data from the target domain with the same masked language modeling objective. Task-adaptive pretraining is similar, but the amount of unlabeled data is smaller and more specific to the task. For instance, when using continued pretraining to enhance model adaptability for progress dataset, DAPT used the entire MIMIC-III dataset, while TAPT only used the training set (without labels) of progress. Previous studies on general domain datasets[53] revealed that both DAPT and TAPT improved cross-domain performance, and the best results were achieved by combining them sequentially (i.e., DAPT+TAPT). Therefore, we conducted experiments with pretrained transformer models adapted using DAPT or DAPT+TAPT. In these

experiments, DAPT, TAPT, or DAPT+TAPT training was performed on top of BioBERT, followed by fine-tuning labeled examples in source domains. We refer to these experiments as DAPT + FT_{source} and DAPT + TAPT + FT_{source}.

It's worth noting that previous pretraining language model work in the clinical domain can be considered as DAPT. For instance, BioClinicalBERT[58] was created by performing DAPT on MIMIC-III using BioBERT[57] as a starting point. Therefore, in this study, we used the existing BioClinicalBERT checkpoint as our DAPT model when progress was the target domain dataset. When thyme was the target domain, we performed continued pretraining for DAPT using an unreleased section of additional unlabeled notes from the patients in the THYME corpus[54]. For discharge, no additional unlabeled data was available, so we used BioClinicalBERT as the DAPT model as a proxy. In DAPT pretraining, we followed the setup described in the BioClinicalBERT paper[58]. It employed a maximum training step count of 15,000 and a learning rate of $5e^{-5}$. For TAPT, we followed the continued pretraining paper[53] and trained the model for 100 epochs, and keep other settings the same. Our TAPT experiments used only the training set of the discharge, progress, and thyme datasets, dropping the labels and using only the unlabeled texts.

3.6 CROSS-DOMAIN SECTION CLASSIFICATION WITH CONTINUED PRETRAINING AND TARGET DOMAIN LABELED DATA

In the previous experiments, we fine-tuned BERT using only source domain labeled data. This is to simulate a real-world situation where we have no annotated data at the target site. We next conduct experiments to simulate a scenario where we have a small amount of labeled data at the target domain, by including a few labeled samples during fine-tuning. We tried adding different amounts of target domain samples, ranging from 10 to 50. We also included DAPT and TAPT as

additional factors. These experiments were referred to as $FT_{source+target}$, $DAPT+FT_{source+target}$, and $DAPT+TAPT+FT_{source+target}$.

3.7 RESULTS

In-domain section classification

In Table 3, the underscored cells of the FT columns displayed the results of in-domain section classification experiments. We found they achieved F1 scores above 0.95 for all three datasets.

Table 3. F1 scores of in-domain and cross-domain experiments, with DAPT and TAPT when applicable. The best F1 score for each combination of source and target domain is in bold.

Source domain (→)	discharge			thyme			progress		
target domain (↓)	FT	DAPT + FT	DAPT + TAPT + FT	FT	DAPT + FT	DAPT + TAPT + FT	FT	DAPT + FT	DAPT + TAPT + FT
discharge	0.972	-	-	0.572	0.6	**0.675**	**0.541**	0.5	0.501
thyme	**0.601**	0.469	0.53	0.99	-	-	**0.646**	0.632	0.544
progress	0.656	0.67	**0.749**	**0.717**	0.58	0.528	0.973	-	-

Cross-domain section classification

In Table 3, the non-underscored cells of the FT columns displayed the results of the cross-domain section experiments. When transitioning from in-domain to cross-domain, F1 scores decreased from above 0.95 range to as low as 0.5.

Cross-domain section classification with continued pretraining

In Table 3, when thyme was the target domain, we found continued pretraining (the DAPT+FT and DAPT+TAPT+FT columns) led to reduced performance. The impact of continued pretraining on progress and discharge was mixed. No substantial performance enhancement was observed when continued pretraining (DAPT or DAPT+TAPT) was directly applied to cross-domain section classification.

Cross-domain section classification with continued pretraining and target domain labeled data

Figure 1 showed the performance change when adding a varying number (10-50) labeled data from the target domain during fine-tuning. It was found that with a larger number of target domain labeled data, the model performance improved. By comparing before and after continued pretraining (DAPT or DAPT+TAPT) was applied, we observed that continued pretraining generally improved model performance, particularly when combined with a small amount of target domain instances.

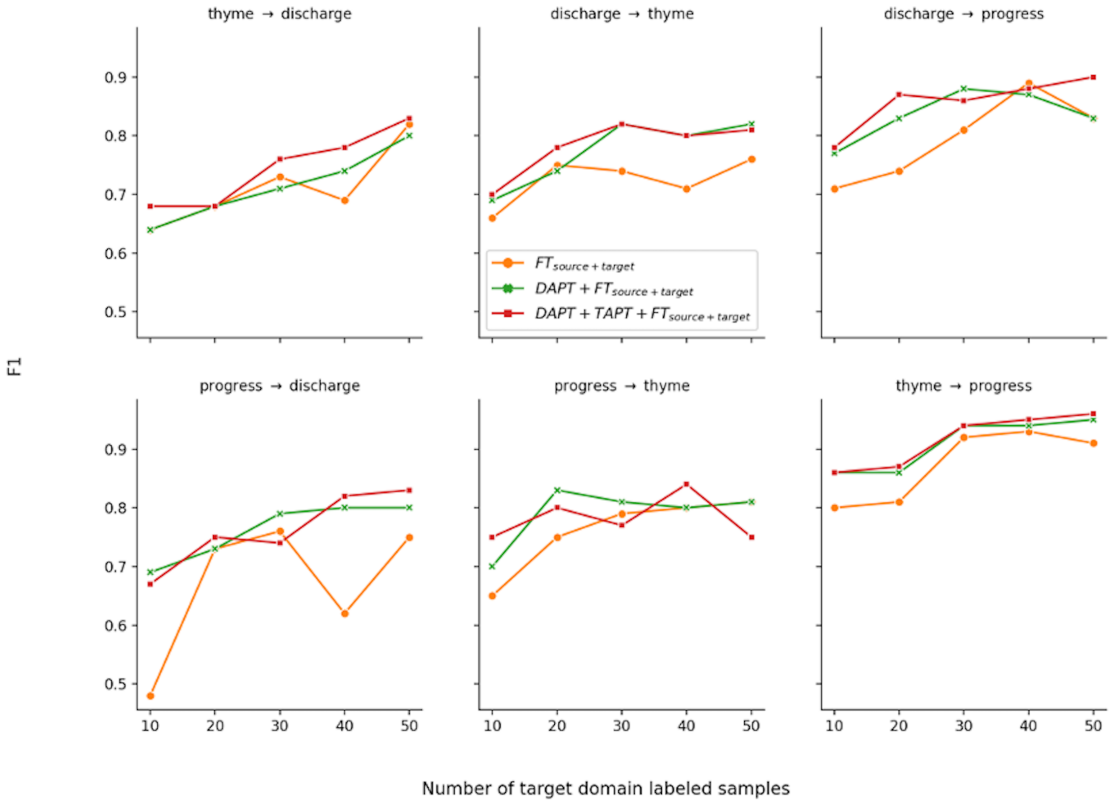

Figure 1. F1 scores of $FT_{source+target}$, $DAPT+FT_{source+target}$, and $DAPT+TAPT+FT_{source+target}$ with 10, 20, 30, 40 and 50 target domain samples for different source and target domain experiments. For example, thyme → discharge represents the experiment with thyme being the source domain and discharge being the target domain.

The F1 scores for $FT_{source+target}$ showed noticeable fluctuations when the source was progress and the target was discharge (see bottom left plot of Figure 1). This is likely because of the noise that can arise with limited labeled target data. This might also be worsened by the significant

difference in section count and section ratio between these two datasets. For discharge, nearly all of its section categories are about ten times fewer than those of progress. When it comes to section ratio, "Assessment" section category makes up 17.7% of the total sections in discharge, while it is only 5.9% in progress. "Others" section category makes up 1.6% of sections in discharge but 7.6% in progress.

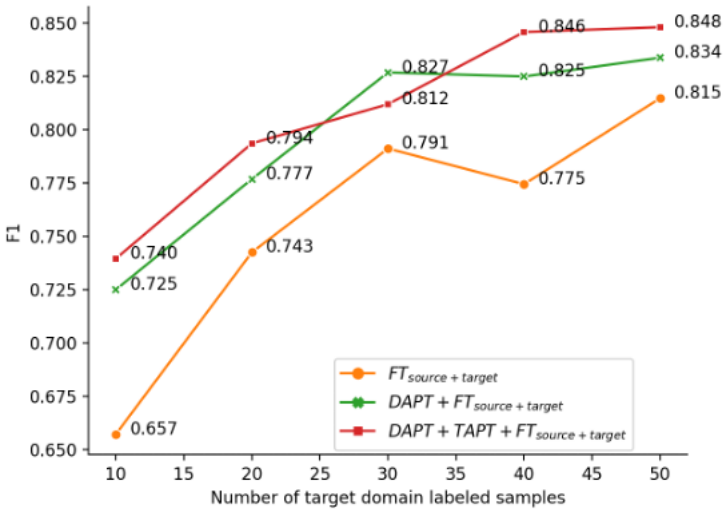

Figure 2. Dataset averaged F1 scores of $FT_{source+target}$, $DAPT+FT_{source+target}$, and $DAPT+TAPT+FT_{source+target}$ with 10, 20, 30, 40 and 50 target domain samples included in fine-tuning.

In Figure 2, we showed the F1 score curves by averaging the six sub-experiments in Figure 1. Notably, $FT_{source+target}$ shows a decline in performance at sample size 40, likely due to the fluctuating progress → discharge model performance observed in Figure 1. On average, the inclusion of continued pretraining (DAPT or DAPT+TAPT) consistently improved the model, when compared to $FT_{source\ target}$. A comparison within the continued pretraining models ($DAPT+FT_{source+target}$ and $DAPT+TAPT+FT_{source+target}$) shows that applying TAPT after DAPT led to a further increase in the F1 score for four out of five sample sizes. Overall, DAPT+TAPT raised the F1 score from 0.756 to 0.808, with an increase of 0.052.

3.8 DISCUSSION

In summary, classifying SOAP sections across different domains is a challenging task, even with task simplifications and the recent modeling advancements. Our experiments with domain adaptation indicate that unsupervised methods alone may not be sufficient but combining them with some annotated data in the new domain can lead to promising results.

Challenges of transferring SOAP section classification

Our findings reveal that while being a straightforward task for humans, current state-of-the-art supervised methods still have struggles generalizing SOAP section classification across domains. This difficulty may arise from variations in documentation practices among different institutions, diverse electronic health record note types, and changing section category distributions. Many tasks are not adequately tested cross-domain, and our study demonstrates that even seemingly simple tasks are hard to generalize. These results echo similar findings in a more detailed version of the task and other clinical NLP tasks[16,60]. Continued pretraining are found to helpful for improving the model performance, but the cross-domain challenges still remain unsolved.

Transferability difference by SOAP category

To determine which SOAP categories are more transferable, we further investigated the cross-domain experiments by breaking down the F1 scores by SOAP categories, and averaging over source-to-targe experiments. The F1 scores by category are as follows: 0.68 for “Subjective,” 0.73 for “Objective,” 0.15 for “Assessment,” 0.13 for “Plan,” and 0.05 for “Others,”.

“Subjective” and “Objective” are the most common categories and exhibit greater transferability compared to the rest. Together, they make up over 70% of the sections. In contrast, the other three categories are less frequent and their prevalence varies significantly

across datasets. For instance, “Plan” appears in 7.5% of discharge dataset, 2.4% in thyme, and not at all in progress.

Additionally, we observed that “Subjective” and “Objective” categories typically have more consistent and similar section names across datasets. Sections like “Chief complaint”, “Social history”, and “Vital sign” are found in all datasets. In contrast, the other three categories often have section names specific to certain datasets, such as “Special instructions”, which is only present in the “Plan” category of thyme.

Chapter 4. ADDRESSING DATA IMBALANCE - CODING HIGHLY RARE CIRCUMSTANCES IN FEMALE FIREARM SUICIDE REPORTS WITH LARGE LANGUAGE MODELS

4.1 OVERVIEW

In this chapter, we employed novel natural language processing methods for coding the rare circumstances in female firearm suicide reports, such as isolation loneliness and bullying (see Table 4 for the full list) [20]. In recent findings, large language models (e.g., ChatGPT) have demonstrated significant power in tasks including yes/no question answering and document classification[19,61]. These models were created from extensive corpora from the web, and they were used to solve machine learning problems in a conversational manner. This study explored adopting a large language model methodology to coding female firearm suicide by designing our coding task as a yes/no question-answering problem and we compared the performance of large language models to the more traditional machine learning approaches.

4.2 DATASET

We utilized the NVDRS Restricted Access Database to study female firearm suicides between 2014 and 2018. This dataset included narrative reports from CME (coroners or medical examiners) and LE (law enforcement), describing the events leading to the suicides of 1,462 females. We manually categorized 9 infrequent circumstances preceding these suicides, such as sleep problems, abusive relationships, custody issues, sexual violence, isolation/loneliness, substance abuse, dementia, bullying, and caregiver issues, following the guidelines in Goldstein et al.[38]. The prevalence of the circumstances can be found in Table 4, and the definitions can be

found in Table 5. Each of these infrequent labels appeared in fewer than 5% of the cases. We divided the dataset into training and test sets with a 50/50 split.

Table 4. Prevalence of the rare female firearm suicide circumstances.

Circumstance	**Count (%)**
Sleep problem	49/1462 (3.4%)
Abusive relationship	44/1462 (3.0%)
Custody	43/1462 (2.9%)
Sexual violence	38/1462 (2.6%)
Isolation loneliness	37/1462 (2.5%)
Substance abuse	32/1462 (2.2%)
Dementia	26/1462 (1.8%)
Bullying	23/1462 (1.6%)
Caregiver	21/1462 (1.4%)

4.3 MODELS

ChatGPT is the state-of-the-art large language model, but due to sensitive data, it cannot be applied directly in our research. Instead, we have turned to open-source alternatives, such as FLAN-T5 and FLAN-UL2[62,63]. Our preliminary experiments showed that FLAN-UL2 was the most effective, and we adopted it for our subsequent experiments. FLAN-UL2 is an open-sourced model developed by Google. It is an encoder-decoder model with 20 billion parameters and performs well question answering. As a baseline comparison, we trained SVM models[64] to identify the rare circumstances as well. FLAN-UL2 was exclusively applied to the test set, while SVMs were trained on the training set and then applied to the test set.

4.4 EVALUATION

In the domain of large language model, a prompt is referred to as the input to the model and it guides a large language model in generating responses. For FLAN-UL2, our prompt consists of a suicide report followed by a question that varies based on the circumstance we want to code. For

instance, when coding "bullying," the question is formulated as follows: "Answer the following yes/no question: was the decedent experiencing bullying in-person or online? Answer:". The model will answers with "Yes" if the circumstance "bullying" is found in the suicide report and "No" if not. This question design is adapted from the circumstance definitions created in the previous study[38]. The adapted questions can be found in Table 5.

Table 5. Circumstance definitions and the questions adapted from them.

Circumstance	**Definition**	**Question**
Sleep problem	The decedent was noted to be experiencing insomnia, sleep associated anxiety, sleep apnea, and/or trouble sleeping.	Was the decedent experiencing insomnia, sleep associated anxiety, sleep apnea, and/or trouble sleeping?
Abusive relationship	The decedent was noted to be providing care for a family member who required long-term assistance due to a chronic illness or physical limitation.	Was the decedent experiencing or recently got out of an intimate relationship that was described as abusive?
Custody issues	The decedent was noted to be having an ongoing legal child custody proceeding, had lost custody of their child, or have lost contact with their child.	Was the decedent having an ongoing legal child custody proceeding, had lost custody of their child, or have lost contact with their child?
Sexual violence	The decedent was noted have experienced sexual violence anytime in their lifetime.	Had the decedent experienced sexual violence anytime in their lifetime?
Isolation loneliness	The decedent was noted to not belong to any social environments or noted to be experiencing loneliness.	Did the decedent not belong to any social environments or was the decedent experiencing loneliness?
Substance abuse	The decedent was noted to have had a history of involvement with a substance use disorder treatment program.	Had the decedent had a history of involvement with a substance use disorder treatment program?
Dementia	The decedent was noted to be experiencing dementia or memory problems including diagnosed Alzheimer's disease, Lewy Body dementia, and Parkinson's disease.	Was the decedent experiencing dementia or memory problems including diagnosed Alzheimer's disease, Lewy Body dementia, and Parkinson's disease?
Bullying	The decedent was noted to be experiencing bullying in-person or online.	Was the decedent experiencing bullying in-person or online?
Caregiver issues	The decedent was noted to be providing care for a family member who required long-term assistance due to a chronic illness or physical limitation.	Was the decedent providing care for a family member who required long-term assistance due to a chronic illness or physical limitation?

We conducted five repetitions of all experiments with training and test set resampling without replacement. The results were reported as the average F1 score with standard deviation.

4.5 Results

We found FLAN-UL2 outperformed SVM in nearly all circumstances, often by a significant margin (see Figure 3). For “sleep problem” and “sexual violence,” the F1 score of FLAN-UL2 exceeded 0.8. The only situation where SVM performed better than FLAN-UL2 was “bullying,” where both SVM and FLAN-UL2 had F1 scores close to zero.

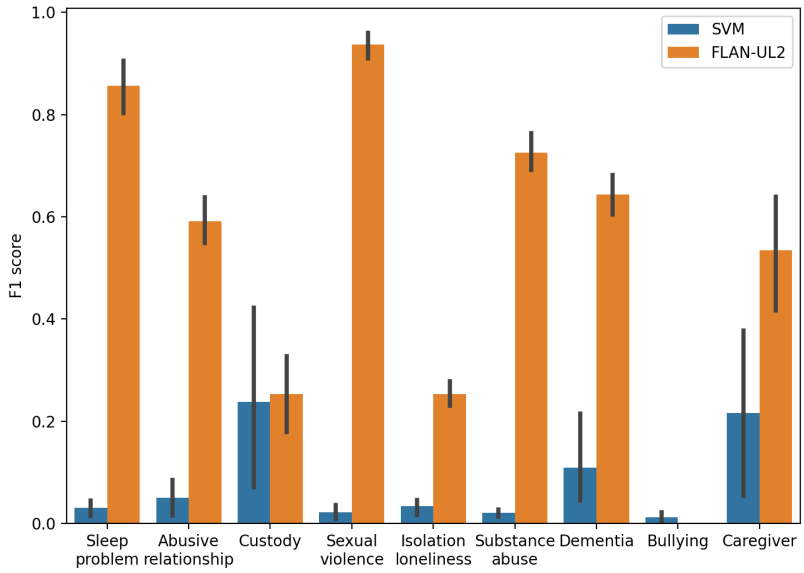

Figure 3. F1 scores of SVM and FLAN-UL2 for coding rare female firearm suicide circumstances from suicide reports when averaged over 5 runs. The height of the bars represents the mean F1 score and the line at the tip of the bars represents the standard deviation across 5 runs.

4.6 Discussion

It is found that large language models are effective in identifying rare circumstances preceding female firearm suicide, surpassing the traditional NLP methods by a large margin. This discovery implies that large language models can facilitate textual analysis in public health research. Furthermore, our straightforward question-answering approach suggests the potential of using large language models for public health surveillance. They could help monitor infrequent conditions that are not previously coded in surveillance systems and are difficult to

code automatically with traditional machine learning methods. Subsequent research should investigate the performance of other emerging large language models and their underlying mechanisms when applied to public health problems.

We discovered that FLAN-UL2 outperformed SVM in identifying 8 out of 9 the rare circumstances preceding female firearm suicide. This implies that using large language models has the potential in filling the gap in detecting rare circumstances in unstructured text efficiently.

Studying the circumstances preceding female firearm suicides is also an overlooked area. A previous study by Goldstein et al.[38] used traditional NLP methods to predict 5 circumstances from suicide reports, with F1-scores between 0.6 and 0.8, but those circumstances had at least 15% prevalence in the dataset. In our study, all 9 infrequent circumstances had less than 5% prevalence. We expanded the previous work by providing a framework for automating the coding of circumstances preceding female firearm suicides on a larger scale.

Error analysis and computational costs

This study found that both SVM and FLAN-UL2 failed in classifying the circumstance "bullying" correctly and it may be because it is one of the most infrequent circumstances. Our question to the large language model, "was the decedent experiencing bullying in-person or online?" might be too vague to answer as well. A more detailed description of "bullying", such as specifying if the victim was insulted or hurt at school or work, may help the model's reasoning. This also suggests that large language models are prompt-sensitive, and designing an appropriate prompt remains crucial for their effectiveness. Additionally, large language models have a high computational cost. Our experiments used 2 NVIDIA A100 40GB GPUs and they might not be affordable for all.

Chapter 5. EXPLORING SYMPTOM TRAJECTORIES PRIOR TO OUT-OF-HOSPITAL CARDIAC ARREST

5.1 OVERVIEW

Out-of-hospital cardiac arrest (OHCA) is a significant public health issue, responsible for 15-20% of global deaths[21]. Despite advancements in resuscitation science and care[65], survival rates for OHCA remain low, with a median survival rate of about 10%[44,45]. To reduce the global impact of cardiac arrest, effective strategies for early recognition and prevention are necessary[46].

Previous research[23–25,47] shows that about half of OHCA patients have symptoms in the hours or days before the event, but only a few seek emergency medical services (EMS) during these warning signs[23]. These studies are useful for those at immediate risk of OHCA, but more information on symptoms over a longer period would help in early risk detection and intervention. A further limitation of current studies[23,24,47] is the reliance on reports from bystanders or emergency medical services (EMS), which can be affected by recall bias. Gathering symptoms reported by patients themselves would offer valuable insights for both individuals and healthcare providers.

In this study, we combine a population-based EMS OHCA registry with longitudinal electronic health records of about 1.5 million patients, including all clinical documentation text. Using a case-control study design with 2,366 OHCA cases and 23,660 controls, we analyze symptom patterns in the year before OHCA using the linked data and natural language processing (NLP) techniques.

5.2 DATASET

The study cohort contains individuals receiving healthcare services at the University of Washington (UW) Medicine healthcare system between January 1, 2010, and December 31, 2021. Situated in Seattle, King County, a metropolitan area with around 2.3 million residents, UW Medicine includes three hospitals and it provided healthcare services for 1.5 million unique patients during the study period. Greater King County and the City of Seattle maintain an extensive registry of all EMS-treated cardiac arrests, documenting arrest rhythm, treatment, and outcomes. The UW Medicine Enterprise Data Warehouse (EDW) collects electronic health records across the UW Medicine system, compiling electronic patient records from over 60 sources, such as lab results, microbiology reports, demographics, diagnoses, prescriptions, and electronic notes. To identify cardiac arrest cases within the UW Medicine system, we cross-referenced all health encounters during the study period with a comprehensive registry of out-of-hospital cardiac arrest (OHCA) cases in King County and the City of Seattle. This matching process considered the person's first and last name and date of birth. Exclusion criteria included individuals under 18 or over 90 years of age, those with a prior history of cardiac arrest, previous cardiopulmonary resuscitation, or having an implantable cardioverter-defibrillator at their initial health encounter. With this, we identified 23,66 OHCA cases. The control group was selected from the rest of the individuals from the UW Medicine system using the same exclusion criteria as the cases. Furthermore, control individuals were excluded if they had any documentation of cardiac arrest, cardiopulmonary resuscitation, or an implantable cardioverter-defibrillator at any time during the study period. This ensured that we excluded individuals who have a cardiac arrest during the lead-in period of the study (January 1, 2010, to December 31, 2012), and those who have in-hospital cardiac arrest or cardiac arrest outside of King County between January 1,

2013 and December 31, 2021. Lastly, controls were matched to cases by age and gender at a 1:10 ratio, resulting in 23,660 controls.

5.3 Ascertainment of covariates

We identified data related to out-of-hospital cardiac arrest from the EMS registry, such as cardiac arrest date and presenting cardiac arrest rhythm. Presenting cardiac arrest rhythm is defined as the earliest rhythm recorded by public access defibrillator application or EMS. Demographic details like age, gender, and race, along with international classification of diseases (ICD) codes and medical notes, were gathered from the UW EDW. Patients with known cardiovascular disease were defined to be those with ICD codes for heart failure, coronary heart disease, and myocardial infarction. In the UW EDW, there are various types of encounters, but the NLP analysis specifically focused on outpatient progress notes, which we considered to provide the most relevant and comprehensive information about patient symptoms.

5.4 Ascertainment of symptoms

We used an existing NLP framework to extract symptoms as events[66]. Each symptom event includes a trigger indicating the symptom's occurrence and a set of attributes (e.g., assertion, anatomy) that describe the symptom. For example, in the sentence "She went to the emergency room last month for her complaint of chest pain," the symptom "chest pain" is represented with the trigger "pain," the assertion "present," and the anatomy "chest" (Figure 4).

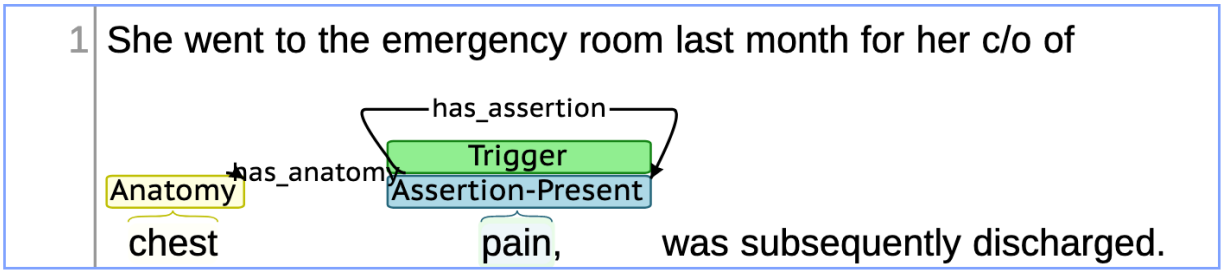

Figure 4. Example of a sentence that contains a symptom event "chest pain". The symptom event is represented with the trigger being "pain", the assertion being "present", and the anatomy being "chest".

To extract symptom events, we employed a previously developed span-based event extraction method with transformer-based language models[7,66]. The model predicts labels for character sequences to identify triggers and various types of attributes and then predicts links between triggers and attributes. We started with a pre-existing language model (BioClinicalBERT) pretrained on ICU patient EHRs[58]. To better fit the health system's notes, we continued pretraining the language model on UW Medicine notes[7,66]. We then trained the model on four datasets annotated with the same symptom event schema. These datasets include the COVID-19 Annotated Text Corpus (CATC), Lung Cancer Annotated Text Corpus (LCATC), Ovarian Cancer Annotated Text Corpus (OCATC), and Post-Cancer Diagnosis Annotated Text Corpus (PCATC)[7,66].

CATC, LCATC, and OCATC contain various types of notes and patients from outpatient clinic patients at the University of Washington Medical Center. CATC covered patients who had clinical visits at the beginning of the COVID pandemic; LCATC contains clinical notes from lung cancer patients but only covered the visits prior to their cancer diagnosis, and OCATC is collected in a similarly way but for the ovarian cancer patients. PCATC is from the Seattle Cancer Care Alliance (SCCA) and includes clinical notes from prostate cancer and lymphoma patients. The combined dataset includes 1028 annotated notes from CATC, 145 from LCATC,

100 from OCATC, and 640 from PCATC. For adaptive pretraining, we collected 87,723 unlabeled notes from the source population of CATC, 281,493 from LCATC, and 29,298 from OCATC. We trained and evaluated the symptom extractor and found it performed at a level comparable to human performance. On CATC, the model's F1 score is 75.4 for triggers and 71.8 for assertions.12, 13 On LCATC, the F1 score is 73.5 for triggers and 67.4 for assertions.12, 13 For OCATC, the F1 score is 79.5 for triggers and 75.1 for assertions[66].

We used this symptom extractor on the progress notes of patients in this study who met specific conditions. We included patients in the symptom analysis if they had at least one progress note available within 720 days of the study end date. For OHCA patients, the study end date is the day of OHCA onset, while for controls, it is the last encounter date. The 720-day window allowed us to focus on notes created close to the OHCA onset but not too close, ensuring most patients in the cohort were included.

The symptom extractor identifies symptoms in free text, which we then categorize for statistical analysis. To evaluate the presence of 13 symptoms related to OHCA, we created a list of synonyms for each symptom (Table 6). We filtered the extracted symptom events to include only those marked as "present" (asserted to be presented by the physician). We then matched these symptoms to our list by first creating an anatomy-trigger string from the extracted symptom event. If this string contains a categorized symptom's synonym, using case-insensitive matching, the extracted symptom is determined to be belong to this categorized symptom.

For example, for a symptom event with the trigger being "had discomfort," assertion being "present" and anatomy being "chest," we created the trigger-anatomy string "had chest discomfort." We then found "chest discomfort," which is a synonym of the categorized symptom "chest pain," was contained in this trigger-anatomy string, we concluded that the patient had the

"chest pain" categorized symptom. For abbreviated synonyms like "SOB," "CP," or "GERD," we required an exact match (case-insensitive) to the extracted symptom event's trigger instead to avoid false positive.

Table 6. The 13 symptoms related to OHCA and their synonyms.

Symptom	**Synonyms**
Shortness of breath	shortness of breath, dyspnea, breathlessness, air hunger, difficulty breathing, rapid breathing, labored breathing, SOB
Diaphoresis	diaphoresis, sweating, hyperhidrosis, perspiration, sudoresis, sweating profusely, sweatiness, clamminess, dampness
Chest pain	chest pain, angina, cardiac pain, chest discomfort, chest heaviness, chest pressure, chest soreness, chest tightness, heartburn, indigestion, CP
Palpitations	palpitation, irregular heartbeat, racing heart, heart racing, skipped beats, heart fluttering, fluttering heart, pounding heart, rapid heartbeat
Light headedness	light headedness, dizziness, dizzy, vertigo, unsteadiness, wooziness, giddiness, spinning sensation
Syncope	syncope, fainting, blackout, collapse, loss of consciousness, pass-out, swoon
Seizure or seizure-like activity	seizure, convulsive syncope, ictal asystole, tonic-clonic movements, myoclonic jerks, automatisms, atonic movements, eyelid fluttering, eye deviation, altered consciousness
Viral respiratory symptoms	fever, cough, sore throat, throat sore, nose running, running nose, nose runny, runny nose, nose stuffy, stuffy nose, muscle aches, joint aches, fatigue, headache, chill, sweat, nausea, vomit
Nausea	nausea, queasiness, stomach upset, upset stomach, stomach discomfort, discomfort stomach, vomit, stomach sick, sick stomach, GERD
Bleeding	bleeding, hemorrhage, blood loss, bleed out, excessive bleeding, spurting blood, oozing blood, profuse bleeding, clotting disorder
Headache	headache, cephalalgia, head pain, painful head, head painful, head pressure, migraine
Peripheral edema	leg swelling, swelling leg, legs swelling, swelling legs, leg swollen, swollen leg, legs swollen, swollen legs, fluid retention, water retention, ankle swelling, swelling ankle, swollen ankle, ankle swollen, leg edema, legs edema, pitting edema, lower extremity edema
Insomnia or sleeping difficulty	sleep problem, difficulty sleeping, insomnia, sleepless, trouble sleeping, wakefulness, inability to sleep, sleep disturbance, restlessness at night, lack of sleep

5.5 STATISTICAL ANALYSIS

To investigate symptoms that impact OHCA over time and for different subgroups, we conduct a statistical analysis between OHCA cases and controls using $\chi 2$ tests or t-tests whenever applicable.

We extracted and compared symptoms over several time windows to observe symptom changes in the year before OHCA. The time periods were defined to expand progressively. All time windows started 720 days (-720) before the study end date (OHCA onset for cases and last encounter for controls). The first window ended 360 days (-360) before the study ended. The second window expanded by 30 days to end 330 days (-330) before the study ended, and this pattern continued until the last window, which ended 30 days before the study ended. This created 12 windows: the first covered 360 days (-720 to -360), the second covered 390 days (-720 to -330), and the last covered 690 days (-720 to -30). Notes from the last 30 days were excluded to avoid including any notes created after OHCA onset. We compared the prevalence of symptoms between cases and controls for each period using $\chi 2$ tests.

We repeated analyses for different groups and variables, including the case/control comparison for male, female, White, non-White, patients with established cardiovascular disease, and patients without established cardiovascular disease. Besides, an analysis on the differences between shockable/non-shockable OHCA was done exclusively among the OHCA patients. We used Bonferroni correction for adjusting to multiple comparisons, with a significance level of $p < 4*10^{-5}$ (12 periods * 13 symptoms * 8 comparisons [total cohort, male, female, White, non-White race, cardiovascular disease, no cardiovascular disease, shockable vs. non-shockable OHCA]). We used multivariable logistic regression to predict and analyze symptoms 30 days before OHCA, including all symptoms and note counts. The model's

accuracy was measured with the area under the ROC curve (AUROC), and precision-recall AUC (AUPRC).

5.6 RESULTS

Study population

In Table 7, throughout the study period, 543,288 clinical progress notes were identified in the study group. This number decreased to 230,393 when the NLP analysis period was limited to 720 days before OHCA for cases and 720 days before the last contact for controls. Out of 2,366 individuals with OHCA, 1,439 (60.8%) had 32,617 clinical progress notes within 720 days of cardiac arrest. Of the 23,660 control patients, 17,238 (72.9%) had 197,776 clinical progress notes within 720 days of the last contact. Baseline characteristics for the analytic cohort are shown in Table 7. Since the case and controls were matched based on gender and age, they had no obvious difference (62±16 vs. 61±16). Women population was around two-thirds for both cases and controls. The case population had a slightly lower proportion of White population than the cases (63% vs. 69%). We found cases patients to be more likely to have established cardiovascular disease (47% vs. 16%). They also had a higher per patient note count (13 vs. 9).

Table 7. Baseline characteristics of the cohort for symptom extraction.

	Out of Hospital Cardiac Arrest (N=1775)	**Controls (N=18,287)**
Age, years	62±16	61±16
Male, n (%)	1146 (65)	11673 (64)
Race, n (%)		
White	1116 (63)	12577 (69)
Non-white	542 (30)	2883 (16)
Unknown	117 (7)	2827 (15)
Established cardiovascular disease, n (%)	896 (38)	3646 (15)
Myocardial infarction	680 (29)	2916 (12)
Heart failure	602 (25)	1631 (7)
Coronary heart disease	387 (16)	1131 (5)
Other Comorbidities, n (%)		
Hypertension	1177 (66)	7442 (41)
Diabetes mellitus	622 (35)	3297 (18)
Hyperlipidemia	1021 (58)	8024 (44)

Atrial fibrillation	346 (19)	1678 (9)
Chronic kidney disease	559 (31)	1826 (10)
Obesity	495 (28)	4106 (22)
Smoking	849 (48)	3904 (21)
Gastrointestinal bleeding	88 (5)	287 (2)
Pulmonary disease	678 (38)	3322 (18)
Obstructive sleep apnea	211 (12)	914 (5)
Stroke or TIA	384 (22)	1501 (8)
Neurological disorders	283 (16)	1083 (6)
Epilepsy	241 (14)	675 (4)
Cirrhosis		
Coagulopathy		
Progress notes per patient, median [IQR]	13 [4,44]	9 [3,27]

Statistics shown are mean ± standard deviation or median [interquartile range], as appropriate.

Symptom trajectories prior to cardiac arrest

In the overall group studied, we found a significant difference in the occurrence of all 13 symptoms between OHCA cases and controls at each time window in the year before the arrest (Figure 5, the values used for creating the figure are in Appendices Table 1). The symptoms with the largest absolute differences between OHCA cases and controls 30 days before the arrest were shortness of breath (36.6% vs. 16.6%), peripheral edema (23.4% vs. 8.8%), viral respiratory symptoms (59.1% vs. 41.2%), and chest pain (28.3% vs. 15.9%). Importantly, the gap in symptom prevalence between OHCA cases and controls widened as the time of cardiac arrest approached.

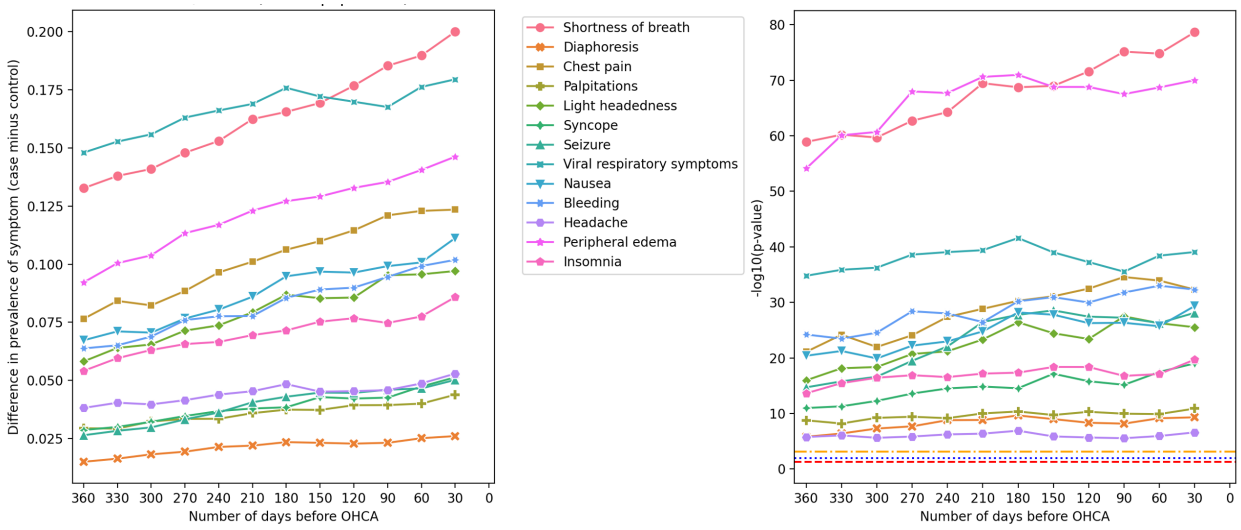

Figure 5. For the overall cohort, symptom trajectories in the year prior to OHCA (case vs. control). Shown are the difference in prevalence of thirteen symptoms between out of hospital cardiac arrest (OHCA) cases compared to controls in the one year prior to OHCA, with each window incrementing by 30 days (Left Panel). The - log_{10} P-values for the comparison of symptom prevalence at each time point are shown (Right Panel). In the right panel, the Bonferroni-adjusted threshold for statistical significance is shown in yellow (yellow-dashed line). For further reference, $p<0.01$ is shown as a blue dashed line and $p<0.05$ is shown as a red dashed line.

Symptoms trajectories in subgroups of interest

In both male and female, shortness of breath, viral respiratory symptoms, and peripheral edema showed the greatest difference between the cases and controls (Figure 6, values for creating the figure in included in Appendices Table 2), having $-\log_{10}$(p-value) > 25 for most of the time windows. However, when looking into the individual male and female groups, we found males having more symptoms to be highly significant. Males have six symptoms viral respiratory symptoms, bleeding, light headedness, chest pain, seizure and nausea to be between the $-\log_{10}$(p-value) of 10 to 30, while females have only two, viral respiratory symptoms and nausea.

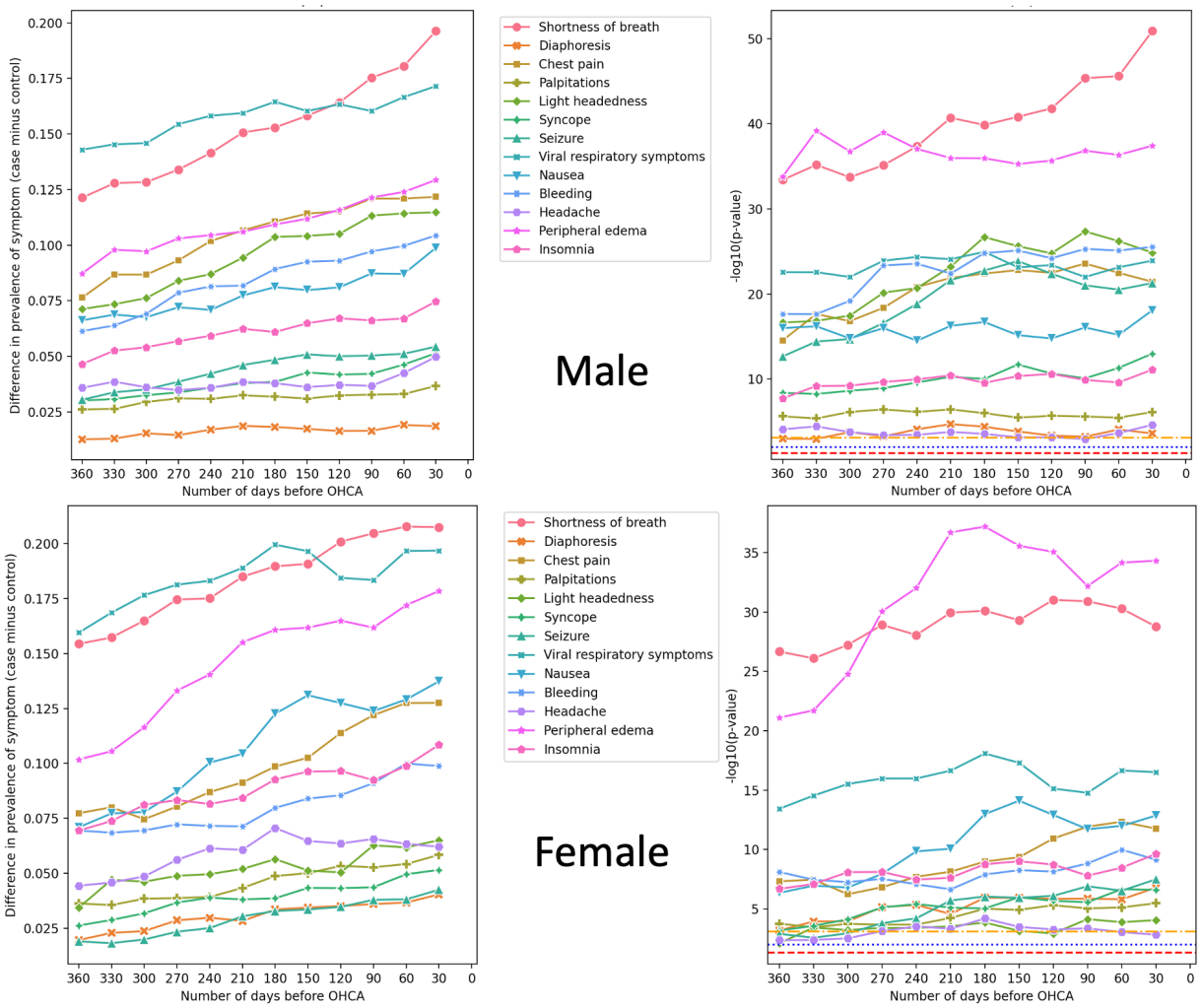

Figure 6. For the male and female cohort, symptom trajectories in the year prior to OHCA (case vs. control). Shown are the difference in prevalence of thirteen symptoms between out of hospital cardiac arrest (OHCA) cases compared to controls in the one year prior to OHCA, with each window incrementing by 30 days (Left Panel). The - log_{10} p-values for the comparison of symptom prevalence at each time point are shown (Right Panel). In the right panel, the Bonferroni-adjusted threshold for statistical significance is shown in yellow (yellow-dashed line). For further reference, $p<0.01$ is shown as a blue dashed line and $p<0.05$ is shown as a red dashed line.

When stratified by race (Figure 7, values for creating the figure in included in Appendices Table 2), symptom enrichment in those of White and non-White race was generally similar to the overall cohort. In the White cohort, we only found one symptom, headache, to have a lower significancy (p > Bonferroni adjusted p-value), and other symptoms all had high significancy (p < Bonferroni adjusted p-value), over almost all time windows. For the non-White cohort, we

found headache and diaphoresis to have a lower significancy (p > Bonferroni adjusted p-value).

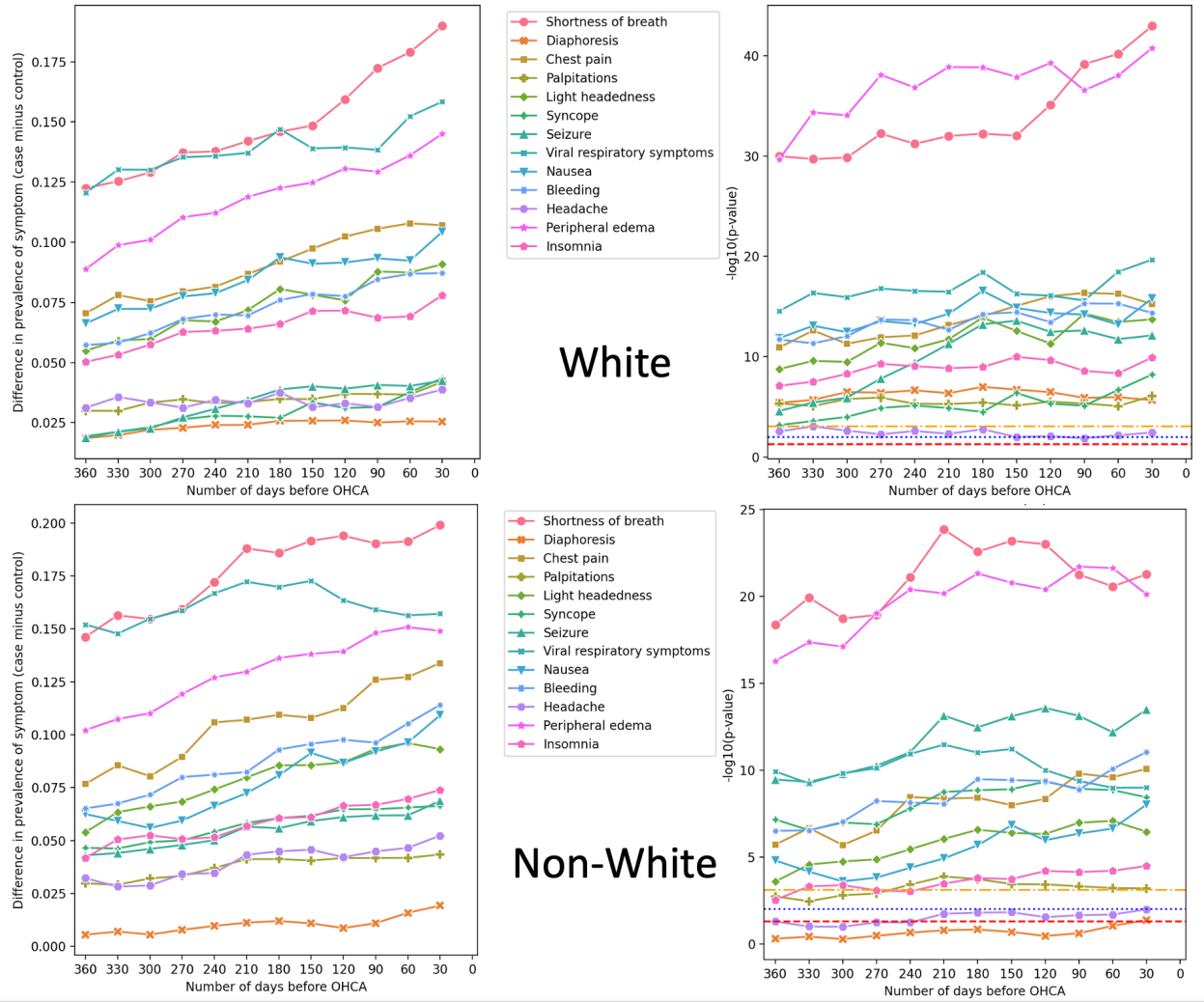

Figure 7. For the White and non-White cohort, symptom trajectories in the year prior to OHCA (case vs. control). Shown are the difference in prevalence of thirteen symptoms between out of hospital cardiac arrest (OHCA) cases compared to controls in the one year prior to OHCA, with each window incrementing by 30 days (Left Panel). The - log_{10} p-values for the comparison of symptom prevalence at each time point are shown (Right Panel). In the right panel, the Bonferroni-adjusted threshold for statistical significance is shown in yellow (yellow-dashed line). For further reference, p<0.01 is shown as a blue dashed line and p<0.05 is shown as a red dashed line.

For individuals with known cardiovascular disease (CVD), the symptoms with -log_{10} (p-value) > 7 are peripheral edema and shortness of breath (Figure 8, values for creating the figure in included in Appendices Table 2). For patients without known CVD, the symptoms are peripheral edema, shortness of breath, viral respiratory symptoms and seizure.

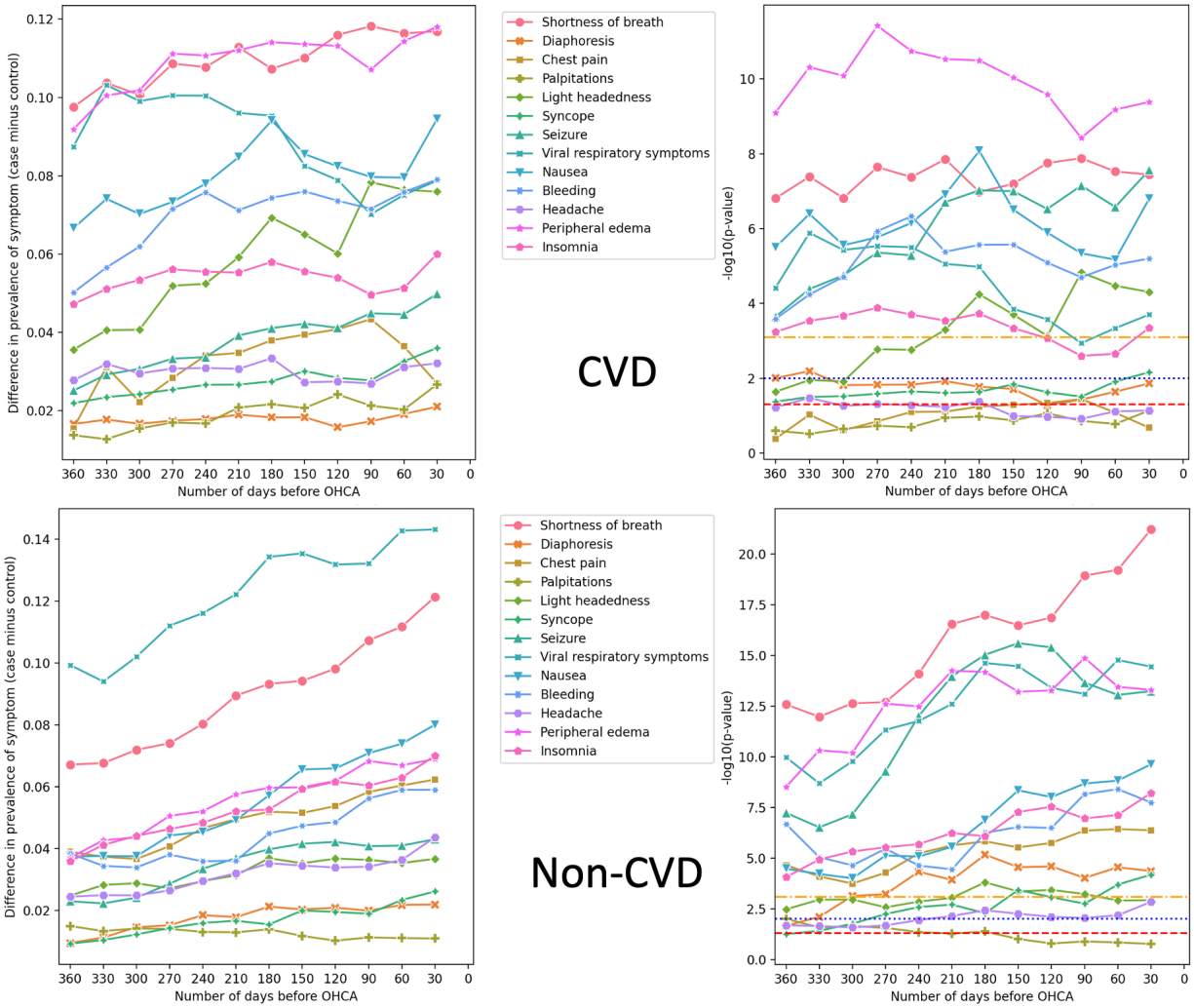

Figure 8. For the CVD and non-CVD (cardiovascular disease) cohort, symptom trajectories in the year prior to OHCA (case vs. control). Shown are the difference in prevalence of thirteen symptoms between out of hospital cardiac arrest (OHCA) cases compared to controls in the one year prior to OHCA, with each window incrementing by 30 days (Left Panel). The - log_{10} p-values for the comparison of symptom prevalence at each time point are shown (Right Panel). In the right panel, the Bonferroni-adjusted threshold for statistical significance is shown in yellow (yellow-dashed line). For further reference, $p<0.01$ is shown as a blue dashed line and $p<0.05$ is shown as a red dashed line.

Symptom trajectories in shockable versus non-shockable OHCA

In an exploratory analysis examining the clinical presentation differences between shockable and non-shockable OHCA, we compared the prevalence of symptoms among OHCA patients. Overall, there were minimal differences in symptoms between the two groups (Figure 9, values for creating the figure in included in Appendices Table 2). Palpitations and chest pain were

slightly more common in shockable OHCA, while other symptoms, including peripheral edema, were more frequently observed in non-shockable OHCA. No symptom showed a significant difference at any time point when comparing the two groups.

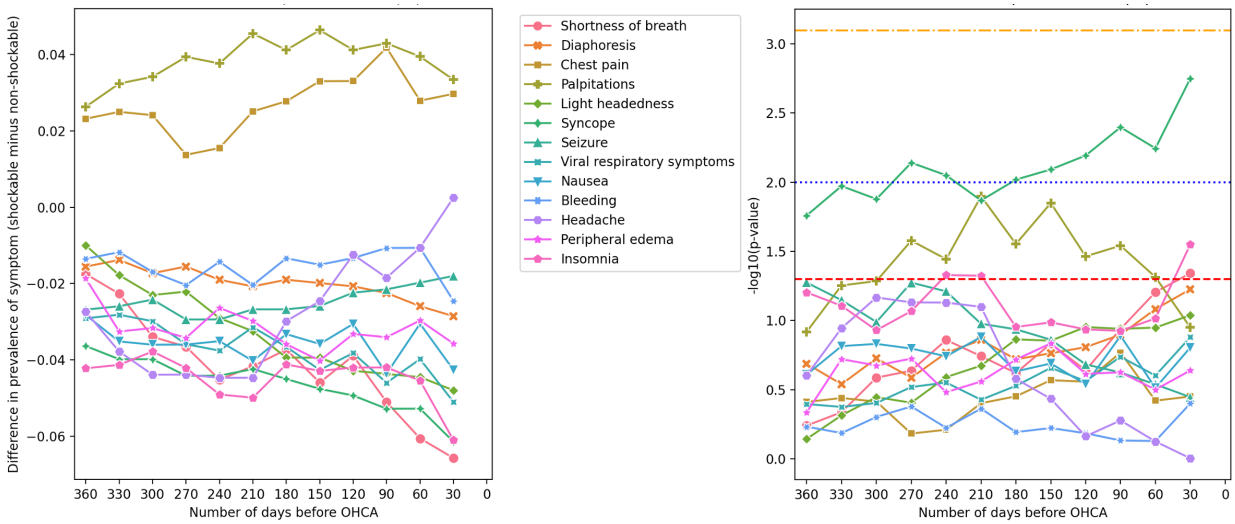

Figure 9. For the OHCA cohort, symptom trajectories in the year prior to OHCA (shockable vs. non-shockable). Shown are the difference in prevalence of thirteen symptoms between shockable hospital cardiac arrest (OHCA) patients compared to non-shockable OHCA patients in the one year prior to OHCA, with each window incrementing by 30 days (Left Panel). The - log_{10} P-values for the comparison of symptom prevalence at each time point are shown (Right Panel). In the right panel, the Bonferroni-adjusted threshold for statistical significance is shown in yellow (yellow-dashed line). For further reference, p<0.01 is shown as a blue dashed line and p<0.05 is shown as a red dashed line.

Multi-symptom models for prediction of out of hospital cardiac arrest

We examined the prediction power and the association between symptoms and out-of-hospital cardiac arrest (OHCA) using a logistic regression model that included all symptoms and adjusted for note density. At 30 days before OHCA, symptoms like peripheral edema (odds ratio [OR]: 1.71, 95% confidence interval [CI]: 1.47-2, p<0.001), shortness of breath (OR 1.85, [95% CI: 1.59-2.15], p<0.001), and seizures (OR 2.21 [95% CI: 1.74-2.81], p<0.001) were significantly linked to higher odds of OHCA (Figure 10, Table 12). The model's ability to discriminate between those who did and did not have an OHCA within 30 days was modest (AUROC = 0.70), as was the precision-recall (AUPRC = 0.17) (Figure 11).

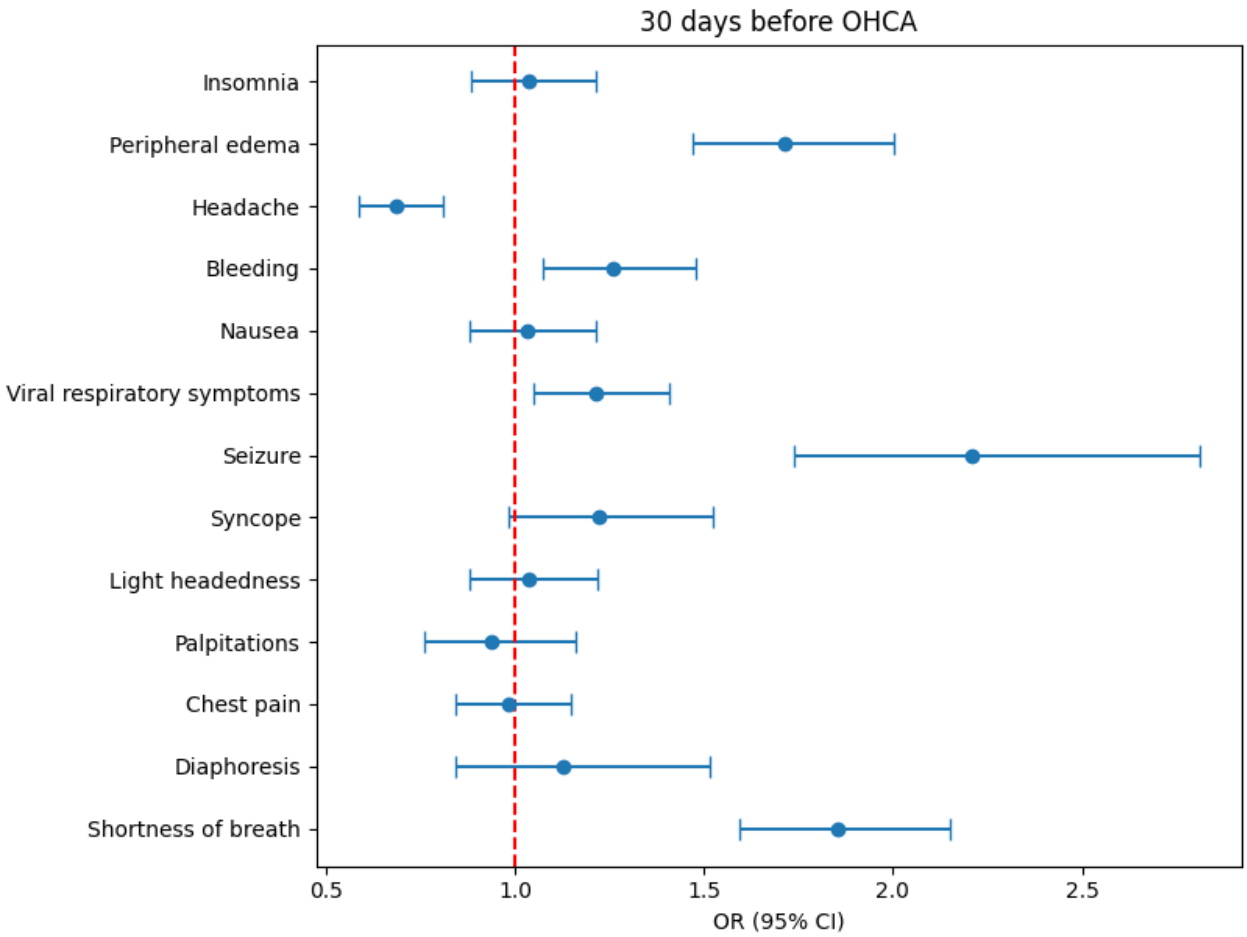

Figure 10. Multivariable association of symptoms in the year prior to OHCA. Shown are the odds (odds ratio [OR] and 95% confidence interval) of out of hospital cardiac arrest (OHCA) associated with symptoms at 30 days prior to cardiac arrest. Multivariable adjustment is for all other symptoms and total number of progress notes present from day 720 to day 30 prior to OHCA.

Table 12. Odds ratio and p-value of multivariable logistic regression Model of symptoms at 30 days prior to OHCA.

Symptom	Odds Ratio (95% Confidence Interval)	P-value
Shortness of breath	1.85 (1.59-2.15)	<0.001
Diaphoresis	1.13 (0.84-1.51)	0.417
Chest pain	0.98 (0.84-1.15)	0.833
Palpitations	0.94 (0.76-1.16)	0.559
Light headedness	1.04 (0.88-1.22)	0.671
Syncope	1.22 (0.98-1.52)	0.073
Seizure	2.21 (1.74-2.81)	<0.001
Viral respiratory symptoms	1.22 (1.05-1.41)	0.009
Nausea	1.03 (0.88-1.21)	0.687
Bleeding	1.26 (1.07-1.48)	0.005
Headache	0.69 (0.58-0.81)	<0.001
Peripheral edema	1.71 (1.47-2)	<0.001
Insomnia	1.04 (0.88-1.22)	0.653
Progress note count	1.01 (1.01-1.01)	<0.001

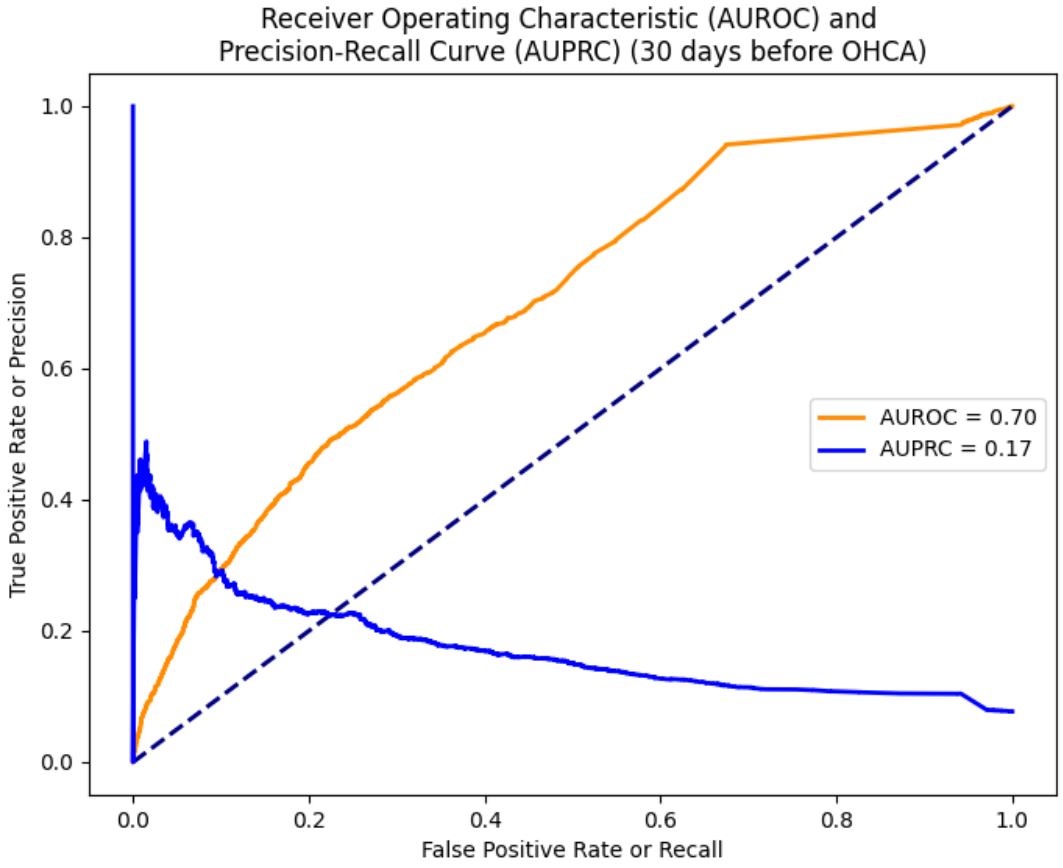

Figure 11. Model performance when predicting OHCA 30 days ahead of time. Shown are the area under the receiver operating characteristic (AUROC) and area under the precision-recall curve (AUPRC) for the association of symptoms at 30 days prior to out of hospital cardiac arrest (OHCA) and the risk of OHCA.

5.7 DISCUSSION

In a general healthcare-seeking population, a wide variety of symptoms – such as those related to the heart and lungs, blood, infections, and nerves – are more common in individuals who experience out-of-hospital cardiac arrest (OHCA) compared to those who do not, up to one year before the event. The frequency of these symptoms increases as the time of the arrest approaches. There were no significant differences in symptoms between those who had arrhythmic (shockable) OHCA and those who had non-shockable OHCA.

Previous research on symptoms before out-of-hospital cardiac arrest (OHCA) has mainly looked at the hours or days leading up to the event, replying on bystanders reports and witnesses, which can be subjective to recall bias. Using natural language processing to analyze clinical notes from a large health system for one year before OHCA, our study addresses previous research limitations by evaluating symptoms over a longer period. This approach provides several significant insights. First, various physical symptoms were observed up to one year before OHCA. About one quarter of patients experienced cardiopulmonary symptoms, such as chest pain, peripheral edema, and shortness of breath, within 30 days of OHCA. Additionally, other notable symptoms found in the overall patient group included viral respiratory symptoms, bleeding, seizure-like activity, and insomnia. These findings indicate the diverse and complex nature of OHCA[65] and support public health efforts to consider both cardiopulmonary and non-cardiopulmonary conditions that increase OHCA risk.

Second, we found that the symptoms in people with shockable OHCA were similar to those with non-shockable OHCA. Traditionally, shockable OHCA is thought to result from cardiovascular issues, while non-shockable OHCA is attributed to non-cardiovascular causes[65].

Recent research shows that non-shockable OHCA is diverse, with many cases linked to heart failure and cardiovascular problems. In our study, although there were more reports of chest pain and palpitations in shockable OHCA, the difference in symptoms between the two groups was not statistically significant at any point in the year before the OHCA. Even among patients with known cardiovascular disease, non-cardiac symptoms like seizure-like activity and insomnia were notably present in the months leading up to OHCA. These results indicate that a wide range of symptoms can precede both types of OHCA, emphasizing the varied nature of the underlying pathologies.

Finally, when considering all symptoms in a multivariable for predicting OHCA 30 days ahead of its onset, the ability to distinguish between OHCA cases and control patients was limited. Our results indicate that symptoms alone are not adequate for screening OHCA risk in the general population. As prediction tools for sudden death improve, we expect that analyzing symptoms through natural language processing could significantly enhance prediction models. At the very least, considering the timeline of symptom development, computational methods like those used in our study could offer patients and healthcare providers an opportunity for early risk assessment and intervention.

This study has several limitations. First, we focused on a pre-selected set of symptoms associated with OHCA. Future studies using a broader and more open-ended approach might find new symptoms and patterns that we did not identify. Second, our natural language processing relied on symptom synonyms and mainly used clinical progress notes. Although our symptom definitions were specific, their sensitivity is unknown, and a wider definition might show different symptom patterns. Including more clinical documents, such as emergency room visits and discharge summaries, could also give more insight into symptom patterns. Third, this

study used electronic health records from a tertiary academic medical center, so the results may not apply to other communities or people not engaged with the healthcare system. Fourth, we recognize that shockable OHCA can change to non-shockable over time, and the classification depends on the time between OHCA onset and rhythm check, possibly causing some misclassification. Lastly, our analysis involved multiple comparisons across different subgroups. Although we used a Bonferroni correction to adjust the significance threshold, there is still a risk of type 1 error.

Chapter 6. HEALTHCARE DISPARITIES AMONG LONG COVID PATIENT POPULATIONS

6.1 OVERVIEW

In this chapter, we seek to understand service access patterns of Long COVID patients from two clinical settings, the primary care clinics and the specialty Long COVID clinic. We took advantage of ICD codes which are recorded as a part of the information in patient visits as structured data. Besides, we used NLP approaches to extract symptoms from clinical notes for gaining insights, covering symptoms that are missing in traditional Long COVID analysis. We also developed machine learning models trained using ICD codes and/or extracted symptoms to predict whether a patient would receive Long COVID specialty care based on their previous clinical records.

6.2 DATASET

To create a comprehensive dataset for this study, we collected electronic health records from patients at the University of Washington Medical Center, spanning from October 2021 to September 2023. For each patient, the collected information included visit information, demographic data, and clinical narratives. We pre-filtered the cohort to include only patients aged 18 and older, having any of the U09.9 Long COVID, U07.1 COVID, and Z86.16 Personal history of COVID in the qualified visits. Visits were collected from the primary care clinics and specialty Long COVID clinic. During data collection, we observed that ICD codes often had creation dates that did not align with their associated visits because physicians sometimes entered these codes before or after the actual visit, which happened when physicians were

planning ahead of the visit or busy moving to the next patient. We remapped these ICDs' code creation dates to their linked visit dates.

Next, we created two study groups for Long COVID patients: the primary care group and the specialty care group. We formed these groups so that we can analyze healthcare allocation among different Long COVID populations. Figure 12 visualized the flow of creating these groups. We first created a cohort containing patients who have visited primary care clinics and the specialty Long COVID clinic, having one of the U09.9/U07.1/Z86.16, having visit records between 10/1/2021 and 9/1/2023, and aged greater or equal than 18. This resulted in 767,254 visits for 9,311 patients. The visits were then split into significant primary care visits and significant specialty care visits, as detailed below.

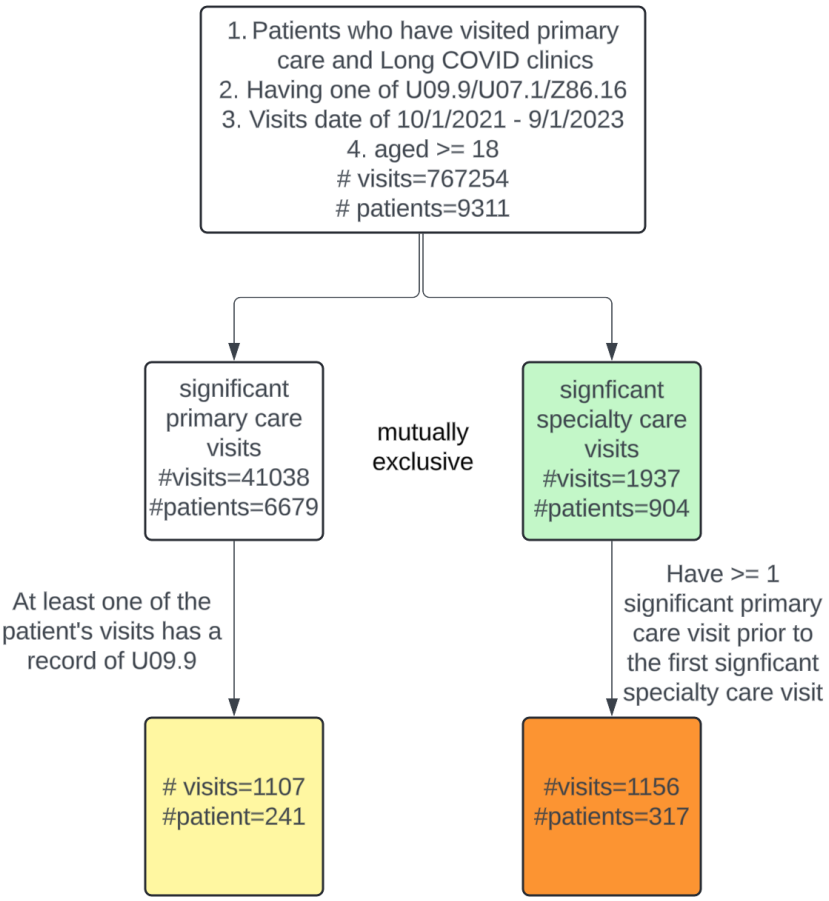

Figure 12. Flow diagram of the cohort selection process for the primary care group, the specialty care group and the retrospective special care group.

The primary care group

To create the Long COVID primary care group, we pre-filtered patient visits to include only those that occurred in primary care clinics (e.g., UWPC Ballard Family Medicine; see Appendix Table 3 for a full list). We selected the substantial visit types (e.g., e-visit, acute visit; see Appendix Table 4 for the full list) and excluded trial visits such as prescription refills and vaccinations. Additionally, the visits were required to be either office visits or telemedicine visits, with a status of completed, thus excluding invalid visits like no-shows. We refer to these as significant primary care visits.

On the basis of the significant primary care visit definition, we further defined the primary care group patients to be those who had at least one U09.9 code, the Long COVID, to be recorded. Furthermore, these patients should not have any specialty care visits, as defined in the next section.

When doing the analysis, we only consider the visits created on/after the first record of U09.9 to avoiding over representation of the symptoms.

The specialty care group

In creating the specialty care group, we first pre-filtered the visits to remove irrelevant ones. Although the specialty Long COVID clinic has a physical location, many Long COVID specialists hosted their visits in any clinics they were based due to the pandemic. In such visits, the physicians were instructed to assign special visit types for these visits: “post COVID new,” “post COVID return,” “telemedicine post COVID new,” and “post COVID telemedicine return.” We filtered the visits to only include these four visit types.

Similar to the primary care group, we also limited the visits to those with a general visit type of either office visit or telemedicine, and a visit status of completed. It is worth noticing that a significant number of patients who visited the special care clinic were referred by a primary care physician and the referrals could be incorrect. In such scenarios, the patient could be in fact non-Long COVID. To confirm that it was a Long COVID visit that actually involved Long COVID care, we further required the visit to have a U09.9 Long COVID code recorded. We refer to these as significant specialty Long COVID visits. Patients who have any occurrences of these visits are considered members of the specialty care group.

The retrospective specialty care group

Among the aforementioned specialty care group patients, there is a subgroup who had a significant primary care visit prior to their first significant specialty care visit. We referred to this subgroup as the retrospective specialty care group.

6.3 GROUPING ICD CODES

The EHR system records thousands of ICD codes, making it difficult to interpret when trying to summarize symptoms that contributed to Long COVID. To facilitate an interpretable analysis, we created ICD groupings based on current practices within the specialty Long COVID clinic and published literatures of Long COVID symptom groupings and a severity of illness index (i.e., the Charlson Comorbidity Index) based on EHR ICD codes.

Long COVID-related diseases/syndromes: Based on current practices in the UW Medicine specialty Long COVID clinic, we identified and proposed disease and syndrome groups that could be related to Long COVID and warrant investigation. We additionally listed the UMLS CUIs (Unified Medical Language System Concept Unique Identifiers) for each group and the CUIs' associated ICD codes, creating a symptom-CUI-ICD code mapping (Appendix Table 5).

We also utilized the symptom-to-ICD mapping in Prado et al.[67] for directly mapping some symptoms to ICDs when applicable.

Psychiatric symptoms: There is evidence of a correlation between psychiatric symptoms and Long COVID[68]. Bobak et al.[68] indicated that preexisting psychiatrics diagnoses, such as anxiety disorders and major depressive disorders , are associated with an increased risk of patients being diagnosed with Long COVID. Based on the ICD-Psychiatric symptom mapping in Bobak et al., we converted our ICD codes to psychiatric symptom groups.

Common comorbidities: To help understand the overall health status of the study groups, we used the symptom-to-ICD mapping defined by the Charlson Comorbidity Index[69] to map our ICD codes and created comorbidities symptom groups (e.g., chronic pulmonary disease, myocardial infarction and diabetes).

6.4 EXTRACTING SYMPTOMS FROM CLINICAL NARRATIVES

Studies have shown that many NLP methods are able to capture a broad range of symptoms from clinical notes[7]. With the goal of creating a more comprehensive representation for the patients, we applied an NLP symptom extractor to the clinical and extracted symptoms. We used the same extractor as described in the OHCA study chapter. Briefly, the extractor extracts symptoms as events. For instance, a sentence "She went to the emergency room last month for her c/o of chest pain, was subsequently discharged," contains the symptom "chest pain," and the extractor will extracted it as a combination of a trigger ("pain"), an assertion ("present"), and an anatomy ("chest"). Different from the OHCA study, here we only used the trigger for mapping synonyms into symptom groups for simplicity.

The NLP algorithm is capable of extracting all symptoms from a note, resulting in a large number of unstructured texts that are difficult to analyze. In an effort of categorizing these

symptoms, we performed a symptom-to-synonym mapping. Based on the symptom-to-synonym mapping table (Appendix Table 6) attached in Lybarger et al.[7], we mapped the extracted symptoms to synonyms using on case-insensitive matching, determined by whether the synonym is a part of the trigger or not. An extracted symptom was categorized into a symptom group when its trigger matched to any synonym from the symptom group.

We applied the NLP extractor to the progress notes, which contain rich information about their potential Long COVID symptoms. The visits we considered were the same as those when used for creating the Long COVID patient study groups.

Lastly, in order to understanding how much the NLP extracted symptoms compliment with the ICD codes, we also created a mapping (Appendix Table 7) that maps the NLP symptoms to each expert derived Long COVID symptoms, which was defined in the previous section, and we showed their distributions among the study groups.

6.5 Predicting patients in need of Long COVID specialty care

The Long COVID clinic provides healthcare to patients who are in need of stepped-up Long COVID care, but the clinic is not resourced to provide full population reach to all patients within UW Medicine. Therefore, it is urgent to develop strategies to identify patients most in need (i.e., of highest priority) to receive access to this care. To address this issue, we developed machine learning models to predict the likelihood that a patient would need specialty care based on their previous primary care visit records.

Patients in the retrospective specialty care group are patients from the specialty care group but had at least one significant primary care visit before they had their first significant specialty care visit. By combining the retrospective specialty care group and the primary care group, we created a patient cohort for training and testing the machine learning model. We took away the

codes U09.9 Long COVID, Z86.16 COVID history and U07.1 COVID from modeling as they were used in the prior study group selection pipeline.

Variables and prediction target

When modeling, we provided all ICD codes and/or grouped NLP extracted symptoms for the primary care group patients. The ICD codes and NLP extracted symptoms were provided as a 0/1 indicator of whether the patient having that ICD code or symptom. For the retrospective specialty care group, we used the ICD codes and/or grouped NLP extracted symptoms created in the significant primary care visits occurring prior to their first Long COVID clinic visit. For the primary care group, we used all of their significant primary care visits. Based on the inputs, the model predicted whether the patient would receive healthcare in the Long COVID clinic.

Machine learning models

We employed multiple machine learning models, including XGBoost, random forest, and neural networks, with hyperparameter search on the training set using 5-fold cross validation and grid search based on AUPRC (Area Under the Precision-Recall Curve). The details of the hyperparameter space for each model are provided in Appendix Table 6. We split the dataset into 50% for training and 50% for testing. By default models used 0.5 as a probability cutoff for determining the label for the binary classification task. We instead used the training set to determine an optimal probability cutoff (i.e., 0.1, 0.2, …., 0.8, 0.9) based on F1 score, and applied it the test set. When applying the trained model on the test set, we repeated the evaluation 100 times by bootstrapping the test set with replacement, and each repetition resulted in metrics such as accuracy, F1 and AUPRC. We reported the average performance metrics, namely accuracy, F1 score, precision, recall, AUROC (Area Under the Receiver Operating Characteristic curve), and AUPRC (Area Under the Precision-Recall Curve).

6.6 RESULTS

Demographics

Table 13 showed the demographics differences between the specialty care and the primary care patient groups. The specialty care group contained 904 patients and the primary care group contains 241, and both of them were 69% female. The specialty care group was 67% White, which was trending higher than the 62% in the primary care group, but was not statistically significantly higher. We observed that the specialty care group patients were younger (age group 35-45, 27% vs. 14%, $p<0.001$) and were covered at higher rates by commercial insurance (commercial insurance 62% vs. 39%, $p<0.001$) than the primary care group.

Table 13. Demographics for the primary care and specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups the for the respective category.

Category	Type	Specialty care	Primary care	Significance
Gender	Male	277 (0.31)	74 (0.31)	n.s.
	Female	627 (0.69)	166 (0.69)	n.s.
	Others	0 (0.0)	1 (0.0)	n.s.
Race	White	610 (0.67)	149 (0.62)	n.s.
	Non-White	294 (0.33)	92 (0.38)	n.s.
Age	18-25	34 (0.04)	**17 (0.07)**	<0.05
	26-35	136 (0.15)	37 (0.15)	n.s.
	36-45	**240 (0.27)**	34 (0.14)	<0.001
	46-55	231 (0.26)	55 (0.23)	n.s.
	56-65	162 (0.18)	43 (0.18)	n.s.
	66-	101 (0.11)	**55 (0.23)**	<0.001
Insurance	Commercial	**559 (0.62)**	94 (0.39)	<0.001
	Non-commercial	345 (0.38)	**147 (0.61)**	<0.001

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Most prevalent ICD codes

Table 14 shows the top 15 most prevalent ICD codes in the specialty care group and their corresponding prevalence in the primary care group, all of which had a prevalence greater than

5% in the specialty care group (i.e., the top 15 ICD codes represent the most prevalent codes documented in specialty care group). The U09.9 Long COVID ICD code was the selection criteria so it had a prevalence of 1.0. The ICD codes that had higher prevalence in the specialty care group compared to primary care group included fatigue (G93.31, G93.3), cognitive issues (R41.89), and insomnia (G47.00). The ICDs that had higher prevalence in the primary care group compared to specialty group included anxiety (F41.9), dyspnea (R06.09) and chronic pain (G89.29). Overall, we identified ICD code usage differences between the two study groups, which can be attributed to either patient differences in actual symptoms or the provider evaluation practices (i.e., providers in primary care may ask about and document symptoms differently than providers in Long COVID specialty care).

Table 14. Top 15 most prevalent ICD codes in the specialty group, and their corresponding prevalence in the primary care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups for the respective category.

Top ICD code	Specialty care	Primary care	Significance
U09.9 Long COVID	904 (1.0)	241 (1.0)	n.s.
G93.31 Postviral fatigue syndrome	**216 (0.24)**	9 (0.04)	<0.001
R41.89 Other symptoms and signs involving cognitive functions and awareness	**199 (0.22)**	12 (0.05)	<0.001
G93.3 Postviral fatigue syndrome	**165 (0.18)**	8 (0.03)	<0.001
R53.83 Other fatigue	156 (0.17)	40 (0.17)	n.s.
R53.82 Chronic fatigue	147 (0.16)	28 (0.12)	n.s.
Z86.16 Personal history of COVID-19	134 (0.15)	29 (0.12)	n.s.
R06.02 Shortness of breath	102 (0.11)	22 (0.09)	n.s.
G93.32 Myalgic encephalomyelitis/chronic fatigue syndrome	100 (0.11)	31 (0.13)	n.s.
G47.00 Insomnia	**85 (0.09)**	12 (0.05)	<0.05
F41.9 Anxiety disorder	81 (0.09)	**35 (0.15)**	<0.05
R51.9 Headache	75 (0.08)	20 (0.08)	n.s.
R06.09 Other forms of dyspnea	68 (0.08)	**31 (0.13)**	<0.05
G89.29 Other chronic pain	67 (0.07)	**70 (0.29)**	<0.001
R42 Dizziness and giddiness	59 (0.07)	17 (0.07)	n.s.

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Table 15 shows the 15 most prevalent ICD codes in the primary care group and their corresponding prevalence in the specialty care group, all of which had a prevalence greater than 10% in the specialty care group. We found the primary care group had 10 out of the 15 ICDs being more prevalent than the specialty care group patients. Given that we used all the visits for the primary care patients that were created after the first record of U09.9, ICD codes for primary care patients might have included codes documented for both regular primary care visits and visited related to Long COVID.

Table 15. Top 15 most prevalent ICD codes in the primary care group, and their corresponding prevalence in the specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups for the respective category.

Top ICD code	Specialty care	Primary care	Significance
U09.9 Long COVID	904 (1.0)	241 (1.0)	n.s.
Z00.00 Encounter for general adult medical examination without abnormal findings	1 (0.0)	**95 (0.39)**	<0.001
Z23 Encounter for immunization	4 (0.0)	**72 (0.3)**	<0.001
G89.29 Other chronic pain	67 (0.07)	**70 (0.29)**	<0.001
I10 Essential (primary) hypertension	8 (0.01)	**60 (0.25)**	<0.001
R53.83 Other fatigue	156 (0.17)	40 (0.17)	n.s.
R05.3 Chronic cough	18 (0.02)	**36 (0.15)**	<0.001
F41.9 Anxiety disorder	81 (0.09)	**35 (0.15)**	<0.05
G93.32 Myalgic encephalomyelitis/chronic fatigue syndrome	100 (0.11)	31 (0.13)	n.s.
R06.09 Other forms of dyspnea	68 (0.08)	**31 (0.13)**	<0.05
M54.50 Low back pain	9 (0.01)	**30 (0.12)**	<0.001
Z86.16 Personal history of COVID-19	134 (0.15)	29 (0.12)	n.s.
R53.82 Chronic fatigue	147 (0.16)	28 (0.12)	n.s.
U07.1 COVID-19	45 (0.05)	**28 (0.12)**	<0.001
E11.9 Type 2 diabetes mellitus without complications	4 (0.0)	**28 (0.12)**	<0.001

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Potential Long COVID symptoms

Table 15 shows the expert derived potential Long COVID symptom distributions for the specialty care group and the primary care group. We found Long COVID-related symptoms, like brain fog (cognitive issues) and difficulty sleeping (insomnia), were more significantly enriched

in the specialty care group. In the contrary, we found general symptoms, like chest pain and diarrhea, were more significantly enriched in the primary care group.

Table 16. Potential Long COVID symptoms grouped by ICD codes distribution for the primary care and specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups the for the respective category.

Potential Long COVID symptom	Specialty care	Primary care	Significance
Fatigue or tiredness	210 (0.23)	49 (0.2)	n.s.
Brain fog	**206 (0.23)**	14 (0.06)	<0.001
Shortness of breath	165 (0.18)	48 (0.2)	n.s.
Difficulty sleeping (too much, too little, early awakening)	**123 (0.14)**	18 (0.07)	<0.05
Dizziness	59 (0.07)	17 (0.07)	n.s.
Malaise	51 (0.06)	6 (0.02)	n.s.
Heart palpitations, pulse skips, heart block	50 (0.06)	14 (0.06)	n.s.
Rapid heart rate (tachycardia)	44 (0.05)	9 (0.04)	n.s.
Memory problems or forgetfulness	36 (0.04)	4 (0.02)	n.s.
Difficulty concentration	27 (0.03)	6 (0.02)	n.s.
Tingling, numbness, burning, stabbing or “pins and needles”	25 (0.03)	**20 (0.08)**	<0.001
Chest pain	18 (0.02)	**16 (0.07)**	<0.001
Diarrhea	18 (0.02)	**14 (0.06)**	<0.01
Indigestion or esophageal/“acid” reflux	15 (0.02)	**25 (0.1)**	<0.001
Nausea	14 (0.02)	6 (0.02)	n.s.
Skin lesions (rash or lumpy lesions)	12 (0.01)	5 (0.02)	n.s.
Loss of smell	11 (0.01)	5 (0.02)	n.s.
Bloating	8 (0.01)	3 (0.01)	n.s.
General/muscle weakness	6 (0.01)	3 (0.01)	n.s.
Abdominal pain	6 (0.01)	**7 (0.03)**	<0.05
Unexplained menstrual irregularity	5 (0.01)	5 (0.02)	n.s.
Unexplained hair loss	5 (0.01)	**12 (0.05)**	<0.001
fever	4 (0.0)	0 (0.0)	n.s.
Muscle pain or cramps	4 (0.0)	3 (0.01)	n.s.
Weight loss	4 (0.0)	**8 (0.03)**	<0.001
Fainting or blackouts	4 (0.0)	**5 (0.02)**	<0.05
Mood swings, irritability, depression	4 (0.0)	**7 (0.03)**	<0.01
Loss of taste	4 (0.0)	1 (0.0)	n.s.
Swelling	3 (0.0)	3 (0.01)	n.s.
Difficulty swallowing	3 (0.0)	**5 (0.02)**	<0.05
Tinnitus (ringing in the ears)	2 (0.0)	1 (0.0)	n.s.
Joint pain or swelling	2 (0.0)	1 (0.0)	n.s.
Problems seeing (double or blurry vision)	2 (0.0)	0 (0.0)	n.s.
Sore throat	1 (0.0)	0 (0.0)	n.s.

Vomiting	1 (0.0)	1 (0.0)	n.s.
Urinary incontinence or difficulty urinating	1 (0.0)	**10 (0.04)**	<0.001
Breathing faster than normal	0 (0.0)	0 (0.0)	n.s.
Cough	0 (0.0)	0 (0.0)	n.s.
Headache	0 (0.0)	0 (0.0)	n.s.
Difficulty hearing	0 (0.0)	0 (0.0)	n.s.
Lack of appetite	0 (0.0)	2 (0.01)	n.s.
Word-finding difficulty	0 (0.0)	0 (0.0)	n.s.
Confusion, difficulty thinking	0 (0.0)	1 (0.0)	n.s.
Sexual dysfunction or loss of libido	0 (0.0)	**4 (0.02)**	<0.01

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Psychiatric symptoms

Table 17 showed the psychiatric symptom prevalence differences between the specialty care and primary care groups. Anxiety disorders were the most prevalent psychiatric symptom in the primary care group. The primary care patients had all psychiatric symptoms more enriched than the specialty care group with statistical significance, except psychotic disorders.

Table 17. Psychiatric symptoms grouped by ICD codes for the primary care and specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups the for the respective category.

Psychiatric symptoms	Specialty care	Primary care	Significance
Anxiety Disorders	111 (0.12)	**65 (0.27)**	<0.001
Other Psychiatric Diagnoses	54 (0.06)	**31 (0.13)**	<0.001
Mood Disorders	49 (0.05)	**26 (0.11)**	<0.01
Major Depressive Disorder	6 (0.01)	**28 (0.12)**	<0.001
Substance Use Disorders	3 (0.0)	**16 (0.07)**	<0.001
Bipolar Disorder	1 (0.0)	**3 (0.01)**	<0.05
Psychotic Disorders	0 (0.0)	1 (0.0)	n.s.

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Charlson Comorbidities Index symptoms

Table 18 showed the Charlson Comorbidities Index symptom prevalence for the two study groups. We observed that the specialty care patients had very few comorbidities documented in ICD codes and the most prevalent symptom, chronic pulmonary disease, only had a prevalence

of 2%. In the contrary, the primary care patients had a higher prevalence for many diseases, such as chronic pulmonary disease, rheumatic disease and mild liver disease, all with statistical significance.

Table 18. Charlson Comorbidities Index (CCI) symptoms for the primary care and specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups the for the respective category.

CCI symptom	Specialty care	Primary care	Significance
Chronic Pulmonary Disease	22 (0.02)	**53 (0.22)**	<0.001
Rheumatic Disease	8 (0.01)	**7 (0.03)**	<0.05
Mild Liver Disease	3 (0.0)	**9 (0.04)**	<0.001
Congestive Heart Failure	3 (0.0)	3 (0.01)	n.s.
Peripheral Vascular Disease	1 (0.0)	2 (0.01)	n.s.
Cerebrovascular Disease	1 (0.0)	**4 (0.02)**	<0.01
AIDS (HIV Infection + opportunistic infection)	1 (0.0)	**7 (0.03)**	<0.001
Hemiplegia or Paraplegia	1 (0.0)	0 (0.0)	n.s.
Myocardial Infarction	0 (0.0)	2 (0.01)	n.s.
Peptic Ulcer Disease	0 (0.0)	0 (0.0)	n.s.
Renal (Mild or Moderate)	0 (0.0)	**5 (0.02)**	<0.001
Diabetes without chronic complication	0 (0.0)	0 (0.0)	n.s.
Any malignancy, including lymphoma and leukemia, except malignant nonmelanoma neoplasm of skin	0 (0.0)	**4 (0.02)**	<0.01
Diabetes with Chronic Complications	0 (0.0)	0 (0.0)	n.s.
Moderate or Severe Liver Disease	0 (0.0)	0 (0.0)	n.s.
Renal (Severe)	0 (0.0)	1 (0.0)	n.s.
HIV Infection	0 (0.0)	0 (0.0)	n.s.
Metastatic Solid Tumor	0 (0.0)	1 (0.0)	n.s.

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

NLP extracted symptoms

Table 19 presented the prevalence of the NLP extracted symptoms for the specialty care and the primary care groups. Some of the symptoms that were significantly more prevalent in the specialty care group included fatigue, headache, anxiety, and shortness of breath. It is worth noticing that NLP captured a lot more symptoms than ICD codes. For instance, in the NLP table, fatigue was found to be recorded in 98% of the specialty care patient's progress note and 67% in the primary care, while in the expert derived Long COVID symptom table, fatigue was less than

50%. In the primary care group, the more prevalent symptoms were cough, diarrhea and swelling.

Table 19. NLP extracted symptoms for the primary care and specialty care group. The ratio for the respective category is shown in parenthesis. A chi-square test is performed between the two groups the for the respective category.

NLP symptoms	Specialty care	Primary care	Significance
fatigue	**855 (0.95)**	171 (0.71)	<0.001
pain	717 (0.79)	184 (0.76)	n.s.
headache	**601 (0.66)**	96 (0.4)	<0.001
anxiety	**526 (0.58)**	104 (0.43)	<0.001
shortness of breath	**523 (0.58)**	121 (0.5)	<0.05
cough	429 (0.47)	**141 (0.59)**	<0.01
fever	**370 (0.41)**	57 (0.24)	<0.001
lightheadedness	**350 (0.39)**	65 (0.27)	<0.001
chills	**275 (0.3)**	37 (0.15)	<0.001
weakness	**273 (0.3)**	55 (0.23)	<0.05
nausea	229 (0.25)	73 (0.3)	n.s.
myalgia	225 (0.25)	56 (0.23)	n.s.
sore throat	**222 (0.25)**	35 (0.15)	<0.01
diarrhea	172 (0.19)	**62 (0.26)**	<0.05
swelling	150 (0.17)	**87 (0.36)**	<0.001
altered mental status	**143 (0.16)**	12 (0.05)	<0.001
ill	128 (0.14)	26 (0.11)	n.s.
gastrointestinal symptoms	**95 (0.11)**	10 (0.04)	<0.01
rash	90 (0.1)	**46 (0.19)**	<0.001
vomiting	88 (0.1)	**39 (0.16)**	<0.01

decreased appetite	76 (0.08)	20 (0.08)	n.s.
dehydration	**74 (0.08)**	6 (0.02)	<0.01
sweats	71 (0.08)	18 (0.07)	n.s.
fall	53 (0.06)	**29 (0.12)**	<0.01
dysphagia	52 (0.06)	9 (0.04)	n.s.
cramping	51 (0.06)	**28 (0.12)**	<0.01
soreness	49 (0.05)	**28 (0.12)**	<0.01
tremors	48 (0.05)	15 (0.06)	n.s.
respiratory symptoms	39 (0.04)	**19 (0.08)**	<0.05
itching	**35 (0.04)**	48 (0.2)	<0.001
wheezing	33 (0.04)	**30 (0.12)**	<0.001
incontinent	30 (0.03)	**24 (0.1)**	<0.001
bleeding	28 (0.03)	**32 (0.13)**	<0.001
heartburn	28 (0.03)	7 (0.03)	n.s.
syncope	27 (0.03)	9 (0.04)	n.s.
runny nose	26 (0.03)	**17 (0.07)**	<0.01
arthralgia	25 (0.03)	**14 (0.06)**	<0.05
irritation	24 (0.03)	**22 (0.09)**	<0.001
erythema	22 (0.02)	**34 (0.14)**	<0.001
ulcers	20 (0.02)	**15 (0.06)**	<0.01
tenderness	18 (0.02)	**75 (0.31)**	<0.001
chest pain	15 (0.02)	0 (0.0)	n.s.
seizures	13 (0.01)	**10 (0.04)**	<0.05
lethargy	10 (0.01)	3 (0.01)	n.s.
bruising	9 (0.01)	**22 (0.09)**	<0.001

sputum	7 (0.01)	**8 (0.03)**	<0.01
discharge	7 (0.01)	**12 (0.05)**	<0.001
wounds	6 (0.01)	5 (0.02)	n.s.
pruritus	6 (0.01)	**9 (0.04)**	<0.001
urinary symptoms	3 (0.0)	2 (0.01)	n.s.
flu-like symptoms	2 (0.0)	1 (0.0)	n.s.
hematochezia	2 (0.0)	3 (0.01)	n.s.
distended	1 (0.0)	**3 (0.01)**	<0.05
deformities	1 (0.0)	**4 (0.02)**	<0.01
exudates	0 (0.0)	2 (0.01)	n.s.

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

In Table 20, the NLP symptoms were mapped to the expert derived Long COVID symptoms and the combined symptom prevalence were shown We found that, for both specialty care patients and primary care clinic patients, adding NLP substantially (with statistically significancy) increased the prevalence of the symptoms of interests, especially for fatigue or tiredness, and headache, where NLP added at least 50% of the prevalence for both study groups. Besides, the differences between the study groups for certain symptoms changed as well. For ICD symptoms (expert derived Long COVID symptoms), fatigue or tiredness had a higher prevalence in the specialty care group than the primary care group but was not significant. After adding NLP symptoms, we it became significant.

Table 20. Potential Long COVID symptoms created from ICD codes, the NLP symptoms mapped to Long COVID symptoms, as well as their combined prevalence. The two inner-group significance value columns (Sig. in the 5th and 9th column) compares the prevalence of ICD and ICD+NLP within a group. The between-group significance column (Sig. in the last column) compares the prevalence of ICD+NLP between the specialty care and the primary care groups.

	Specialty care				Primary care				Sig.
Potential Long COVID symptom	**ICD**	**NLP**	**ICD+NLP**	**Sig.**	**ICD**	**NLP**	**ICD+NLP**	**Sig.**	
Fatigue or tiredness	210 (0.23)	865 (0.96)	**866 (0.96)**	<0.001	49 (0.2)	178 (0.74)	178 (0.74)	<0.001	<0.001
Headache	0 (0.0)	601 (0.66)	**601 (0.66)**	<0.001	0 (0.0)	96 (0.4)	96 (0.4)	<0.001	<0.001
Difficulty sleeping (too much, too little, early awakening)	123 (0.14)	530 (0.59)	**552 (0.61)**	<0.001	18 (0.07)	106 (0.44)	112 (0.46)	<0.001	<0.001
Shortness of breath	165 (0.18)	530 (0.59)	540 (0.6)	<0.001	48 (0.2)	129 (0.54)	131 (0.54)	<0.001	n.s.
Mood swings, irritability, depression	4 (0.0)	530 (0.59)	**533 (0.59)**	<0.001	7 (0.03)	106 (0.44)	108 (0.45)	<0.001	<0.001
Breathing faster than normal	0 (0.0)	523 (0.58)	**523 (0.58)**	<0.001	0 (0.0)	121 (0.5)	121 (0.5)	<0.001	<0.05
Cough	0 (0.0)	429 (0.47)	429 (0.47)	<0.001	0 (0.0)	141 (0.59)	**141 (0.59)**	<0.001	<0.01
Fever	4 (0.0)	370 (0.41)	**370 (0.41)**	<0.001	0 (0.0)	57 (0.24)	57 (0.24)	<0.001	<0.001
Dizziness	59 (0.07)	350 (0.39)	**358 (0.4)**	<0.001	17 (0.07)	65 (0.27)	68 (0.28)	<0.001	<0.01
Brain fog	206 (0.23)	143 (0.16)	**311 (0.34)**	<0.001	14 (0.06)	12 (0.05)	26 (0.11)	n.s.	<0.001
General/muscle weakness	6 (0.01)	273 (0.3)	**274 (0.3)**	<0.001	3 (0.01)	55 (0.23)	55 (0.23)	<0.001	<0.05
Muscle pain or cramps	4 (0.0)	257 (0.28)	259 (0.29)	<0.001	3 (0.01)	72 (0.3)	72 (0.3)	<0.001	n.s.
Nausea	14 (0.02)	229 (0.25)	230 (0.25)	<0.001	6 (0.02)	73 (0.3)	73 (0.3)	<0.001	n.s.
Sore throat	1 (0.0)	222 (0.25)	**222 (0.25)**	<0.001	0 (0.0)	35 (0.15)	35 (0.15)	<0.001	<0.01
Diarrhea	18 (0.02)	172 (0.19)	172 (0.19)	<0.001	14 (0.06)	62 (0.26)	**62 (0.26)**	<0.001	<0.05
Memory problems or forgetfulness	36 (0.04)	143 (0.16)	**170 (0.19)**	<0.001	4 (0.02)	12 (0.05)	15 (0.06)	<0.05	<0.001
Malaise	51 (0.06)	128 (0.14)	170 (0.19)	<0.001	6 (0.02)	26 (0.11)	32 (0.13)	<0.001	n.s.
Joint pain or swelling	2 (0.0)	167 (0.18)	168 (0.19)	<0.001	1 (0.0)	93 (0.39)	**94 (0.39)**	<0.001	<0.001
difficulty concentration	27 (0.03)	143 (0.16)	**161 (0.18)**	<0.001	6 (0.02)	12 (0.05)	18 (0.07)	<0.05	<0.001
Swelling	3 (0.0)	150 (0.17)	150 (0.17)	<0.001	3 (0.01)	87 (0.36)	**87 (0.36)**	<0.001	<0.001
Confusion, difficulty thinking	0 (0.0)	143 (0.16)	**143 (0.16)**	<0.001	1 (0.0)	12 (0.05)	13 (0.05)	<0.01	<0.001
Skin lesions (rash or lumpy lesions)	12 (0.01)	113 (0.12)	113 (0.12)	<0.001	5 (0.02)	70 (0.29)	**71 (0.29)**	<0.001	<0.001
Bloating	8 (0.01)	95 (0.11)	**101 (0.11)**	<0.001	3 (0.01)	10 (0.04)	13 (0.05)	<0.05	<0.05
Abdominal pain	6 (0.01)	95 (0.11)	**99 (0.11)**	<0.001	7 (0.03)	10 (0.04)	15 (0.06)	n.s.	<0.05
Vomiting	1 (0.0)	88 (0.1)	88 (0.1)	<0.001	1 (0.0)	39 (0.16)	**39 (0.16)**	<0.001	<0.01
Lack of appetite	0 (0.0)	76 (0.08)	76 (0.08)	<0.001	2 (0.01)	20 (0.08)	20 (0.08)	<0.001	n.s.
Rapid heart rate (tachycardia)	44 (0.05)	28 (0.03)	70 (0.08)	<0.05	9 (0.04)	7 (0.03)	14 (0.06)	n.s.	n.s.
Difficulty swallowing	3 (0.0)	52 (0.06)	53 (0.06)	<0.001	5 (0.02)	9 (0.04)	10 (0.04)	n.s.	n.s.

Heart palpitations, pulse skips, heart block	50 (0.06)	/	50 (0.06)	n.s.	14 (0.06)	/	14 (0.06)	n.s.	n.s.
Indigestion or esophageal/"acid" reflux	15 (0.02)	28 (0.03)	40 (0.04)	<0.01	25 (0.1)	7 (0.03)	**29 (0.12)**	n.s.	<0.001
Chest pain	18 (0.02)	15 (0.02)	32 (0.04)	n.s.	16 (0.07)	0 (0.0)	16 (0.07)	n.s.	n.s.
Urinary incontinence or difficulty urinating	1 (0.0)	30 (0.03)	31 (0.03)	<0.001	10 (0.04)	24 (0.1)	**32 (0.13)**	<0.001	<0.001
Fainting or blackouts	4 (0.0)	27 (0.03)	29 (0.03)	<0.001	5 (0.02)	9 (0.04)	12 (0.05)	n.s.	n.s.
Tingling, numbness, burning, stabbing or "pins and needles"	25 (0.03)	/	25 (0.03)	n.s.	20 (0.08)	/	**20 (0.08)**	n.s.	<0.001
Loss of smell	11 (0.01)	/	11 (0.01)	n.s.	5 (0.02)	/	5 (0.02)	n.s.	n.s.
Unexplained hair loss	5 (0.01)	/	5 (0.01)	n.s.	12 (0.05)	/	**12 (0.05)**	n.s.	<0.001
Unexplained menstrual irregularity	5 (0.01)	/	5 (0.01)	n.s.	5 (0.02)	/	5 (0.02)	n.s.	n.s.
Weight loss	4 (0.0)	/	4 (0.0)	n.s.	8 (0.03)	/	**8 (0.03)**	n.s.	<0.001
Loss of taste	4 (0.0)	/	4 (0.0)	n.s.	1 (0.0)	/	1 (0.0)	n.s.	n.s.
Problems seeing (double or blurry vision)	2 (0.0)	/	2 (0.0)	n.s.	0 (0.0)	/	0 (0.0)	n.s.	n.s.
Tinnitus (ringing in the ears)	2 (0.0)	/	2 (0.0)	n.s.	1 (0.0)	/	1 (0.0)	n.s.	n.s.
Difficulty hearing	0 (0.0)	/	0 (0.0)	n.s.	0 (0.0)	/	0 (0.0)	n.s.	n.s.
Word-finding difficulty	0 (0.0)	/	0 (0.0)	n.s.	0 (0.0)	/	0 (0.0)	n.s.	n.s.
Sexual dysfunction or loss of libido	0 (0.0)	/	0 (0.0)	n.s.	4 (0.02)	/	**4 (0.02)**	n.s.	<0.01

Note: n.s. = not significant; <x represents p-values smaller than x; bolded values represent the higher value among the two groups and with statistically significance.

Predicting patients in need of Long COVID specialty care

We performed two sets of experiments, one using patients' ICD codes as the model input and another one using both ICD codes and NLP extracted symptoms. The models performance metrics, such as accuracy and F1 score, were shown in Table 19. Overall, the models trained with ICD codes and NLP symptoms resulted in better performance, including a higher AUROC, AUPRC, F1 and recall, showing the potential of using clinical notes to better identify Long COVID patients in need of healthcare. Besides, we found the ICD + NLP symptom models

performed better with statistical significance in most of the cases. The logistic regression model trained with ICD and NLP symptoms achieved the highest F1 score, AUROC and AUPRC. We applied this model on the test set and perform error analysis.

Table 21. The performance metrics for different models used for predicting whether a patient would receive specialty care. The top half of the table used ICD codes for building the model and the bottom used ICD codes and symptoms extracted from progress notes. The numbers shown are the mean by bootstrapping the test set 100 times with replacement, with the standard deviation shown in parenthesis. The bolded number indicates statistical significancy using t-test with p-value < 0.05, comparing between models using ICD and ICD + NLP symptoms.

	Model	***Accuracy***	***Precision***	***Recall***	***F1***	***AUROC***	***AUPRC***
ICD	Neural network	0.65 (0.03)	0.69 (0.04)	0.69 (0.04)	0.69 (0.03)	0.73 (0.03)	0.77 (0.03)
	KNN	0.56 (0.03)	0.57 (0.03)	0.98 (0.01)	0.72 (0.02)	0.71 (0.03)	0.77 (0.04)
	Random forest	0.67 (0.03)	0.71 (0.04)	**0.71 (0.04)**	0.71 (0.03)	0.76 (0.03)	0.83 (0.03)
	Logistic regression	0.57 (0.03)	0.57 (0.03)	**1.0 (0.0)**	0.72 (0.02)	0.78 (0.03)	0.84 (0.03)
	XGBoost	0.7 (0.03)	0.72 (0.04)	**0.76 (0.04)**	0.74 (0.03)	0.78 (0.03)	0.83 (0.03)
ICD + NLP symptom	***Model***	***Accuracy***	***Precision***	***Recall***	***F1***	***AUROC***	***AUPRC***
	Neural network	**0.66 (0.03)**	0.68 (0.03)	**0.76 (0.03)**	**0.72 (0.03)**	**0.74 (0.03)**	**0.79 (0.03)**
	KNN	0.57 (0.03)	0.57 (0.03)	**1.0 (0.0)**	0.72 (0.02)	0.71 (0.03)	**0.78 (0.04)**
	Random forest	**0.72 (0.03)**	**0.8 (0.03)**	0.67 (0.04)	**0.73 (0.03)**	**0.8 (0.02)**	**0.86 (0.02)**
	Logistic regression	**0.64 (0.03)**	**0.62 (0.03)**	0.94 (0.02)	**0.74 (0.02)**	**0.82 (0.02)**	**0.87 (0.02)**
	XGBoost	**0.72 (0.03)**	**0.76 (0.04)**	0.73 (0.03)	0.74 (0.03)	**0.79 (0.03)**	**0.85 (0.02)**

Table 22 showed the confusion matrix for predictions made by the logistic model. We observed that 91 of the 240 (38%) predicted positives were actually negatives, i.e., the model determined that they should be sent to the specialty care clinic but they were not. Figure 14 shows a force plot of such a false positive patient. In the figure, we observed that the patient had fatigue, headache, and Postviral fatigue (G93.3, G93.31). These symptoms were pushing the model to predict the patient to positive.

Table 22. Confusion matrix for the trained logistic regression model when applied on the full test set for predicting whether a patient would be seen in the Long COVID specialty care clinic.

Positives are patients in the retrospective specialty care group and negatives are patients in the primary care group.

	Predicted positive	Predicted negative
Actual positive	149	10
Actual negative	91	30

In the confusion matrix, we also observed that 10 out of the 40 (25%) predicted negative patients were actually positive patients, i.e., the model determined that they should not be sent to the specialty care clinic but they were. Figure 15 showed one such patient's force plot. We observed that the patient had few Long COVID related symptoms, and the patient had Z00.00 Encounter for general adult medical examination without abnormal findings, Z23 Encounter for immunization, and G89.29 Other chronic pain. In the feature importance plot, the model considered these features as indicative that the patient should stay in the primary care clinic. A possible explanation is that the primary care clinicians did not document the Long COVID symptoms fully.

Figure 13 showed the probability density plot for the true positives, false positives, true negatives, and false negatives. We observed that all four kinds of predictions had a probability density centered around 0.5, and that the model tended not to make predictions with probabilities lower than 0.4 and higher than 0.8. The true positives have the most plateaued distribution, with the probabilities distributed more evenly than the others. The false negatives have the most peaked distribution, with most of the probability density centering around 0.5. When comparing the true positives and the false negatives, we found that most of the false positives had probabilities around 0.5, meaning that the model was not overly confident about the type I errors it made.

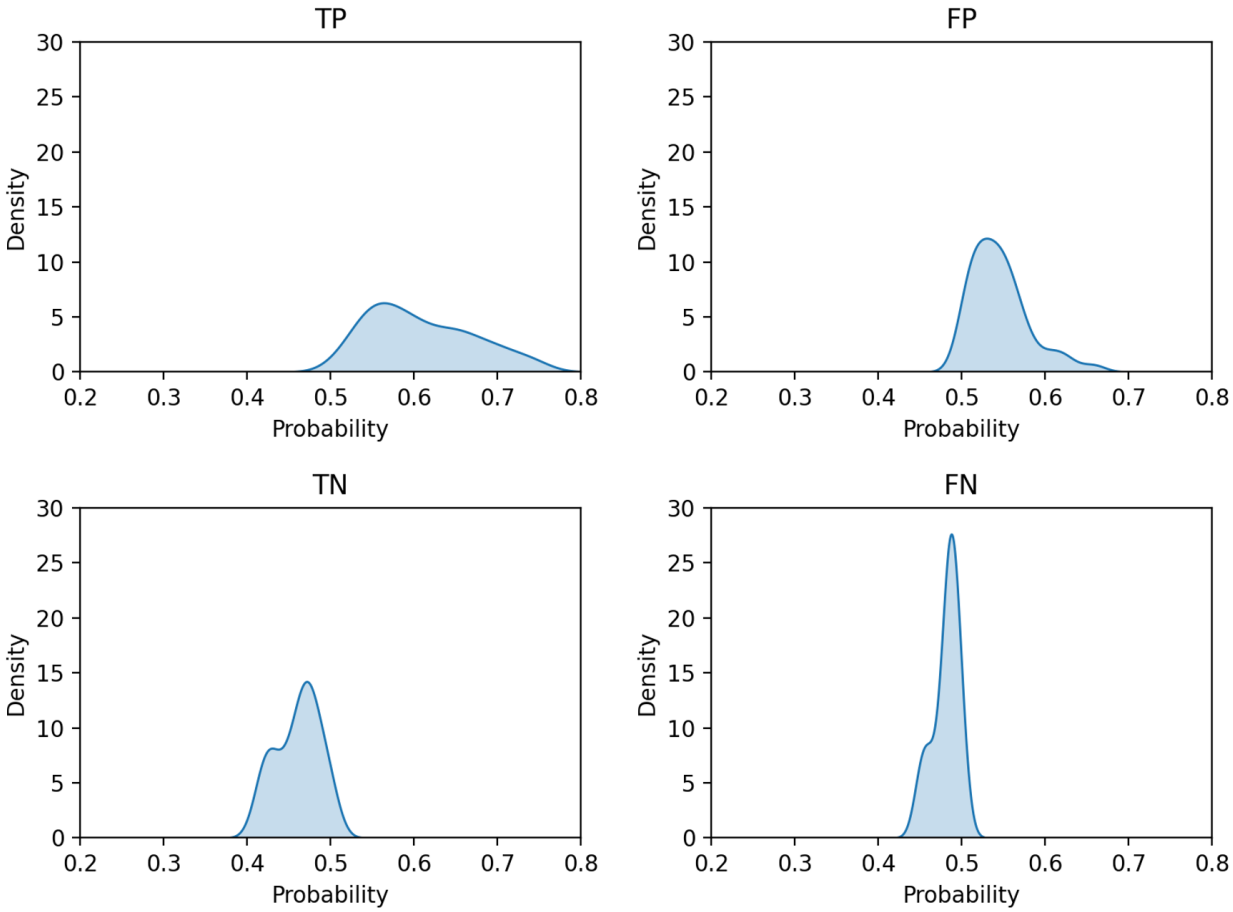

Figure 13. A probability density plot for the prediction made by the logistic regression model. The x axis is probability, the y-axis is the number of predictions with that probability. To help with interpretability, we applied a Gaussian kernel and created a continuous probability density curve. There are four subplots representing different types of predictions, TP (true positive), FP (false positive), TN (true negative), and FN (false negative). The optimal positive/negative prediction probability cutoff was 0.5. The figure had positives/negatives crossing 0.5 because of the Gaussian kernel's smoothing effect.

The errors could be from multiple sources, with one of them being the number of ICD codes/NLP symptoms each patient has. In Table 23, we observed that patients in the positive group (retrospective specialty care group) had a lower number of ICD codes/NLP symptoms than the patients in the negative group (retrospective specialty care group). This was reflected in the different types of predictions too. The patients who were correctly predicted as positive (TP) had a lower number of ICD codes/NLP symptoms than the patients who were wrongly predicted

as positive (FP). The patients who were correctly predicted as negative (TN) had a higher number of ICD codes/NLP symptoms than the patients who were wrongly predicted as negative (FN). Overall, the number of ICD codes/NLP symptoms each patient had might be misleading to the model and causing misclassification.

Table 23. The average number of ICD codes/NLP symptoms each patient had in the test set, stratified by positive (retrospective specialty care group), negative (primary care group), TP (true positive), FP (false positive), TN (true negative), and FN (false negative).

	Positive	Negative	TP	FP	TN	FN
Average number of ICD codes/NLP symptoms for each patient	17.8	22.6	17.1	19.7	31.4	29.3

Another source of the error could be coming from the specific type of ICD codes/NLP symptoms each patient had. Figure 14 showed the top 10 features that the greatest impact on the logistic regression model's prediction - whether a patient would be admitted to the Long COVID specialty care clinic. The feature importance values were calculated based on SHAP values[70]. We found Long COVID related ICD codes/symptoms, namely fatigue, headache, G93.3 Postviral and related fatigue syndromes, and G93.31 Postviral fatigue syndrome were strong indicators for patients being admitted to the specialty care (positive). We found patients with common primary care clinic ICD codes/symptoms, namely Z00.00 Encounter for general adult medical examination without abnormal findings, swelling, and pain were strong indicators of patients staying in the primary care clinics (negative).

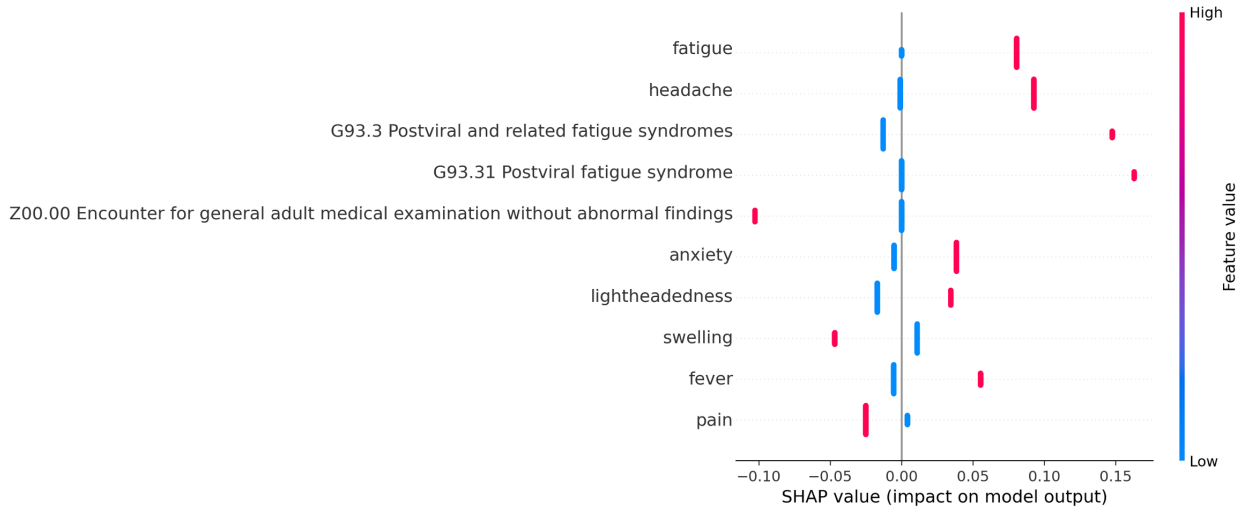

Figure 14. The 10 features that have the greatest impact (SHAP values) on the model prediction. The colored bar represented the presence (red) and absence (blue) of a feature. The x-axis indicated the impact of a feature on model output. Positive SHAP values tended to make the model output positive predictions and negative SHAP values tended to make the model output negative predictions. For instance, when a patient had G93.3 Postviral and related fatigue syndromes (third feature), the model tended to strongly classify the patient into positives; when a patient did not have it, the model tended to mildly classify the patient into negatives.

Table 24 showed the average number of ICD code/NLP symptoms each patient had in the four types of predictions, namely, TP, FP, TN, and FN. In comparing between positive predictions that were correct (TP) and positive predictions that were wrong (FP), we found TP tended to have a higher prevalence for symptoms that had strong indicators of patients being positive figure, e.g., fatigue, headache, G93.3 Postviral and related fatigue syndromes, and G93.31 Postviral fatigue syndrome.

Table 24. The top 10 features that had the greatest impact on model prediction, and their prevalence in different types of predictions.

	TP	FP	TN	FN
fatigue	0.91	0.71	0.77	0.5
headache	0.62	0.38	0.37	0.2
G93.3 Postviral and related fatigue syndromes	0.32	0.04	0	0
G93.31 Postviral fatigue syndrome	0.34	0.04	0	0
Z00.00 Encounter for general adult medical examination without abnormal findings	0.15	0.33	0.67	0.7
anxiety	0.62	0.48	0.43	0.6

lightheadedness	0.42	0.24	0.3	0.2
swelling	0.2	0.32	0.53	0.6
fever	0.34	0.32	0.07	0.1
pain	0.76	0.7	0.83	1

In comparing between negative predictions that were correct (TN) and negative predictions that were wrong (FN), we found the TN tended to have a higher prevalence for symptoms that had strong indicators of patients being negatives, e.g., Z00.00 Encounter for general adult medical examination without abnormal findings, swelling, pain.

We provided two false positive and false negative examples to help understand the prediction errors. In Figure 15, the patient was wrongly classified in to positive, for it having strong positive indicators, namely fatigue, headache, and Postviral fatigue (G93.3, G93.31), although the patient also had Z00.00 Encounter for general adult medical examination without abnormal findings. In Figure 16, the patient was wrongly classified in to negative, for it having strong negative indicators, namely Z00.00 Encounter for general adult medical examination without abnormal findings.

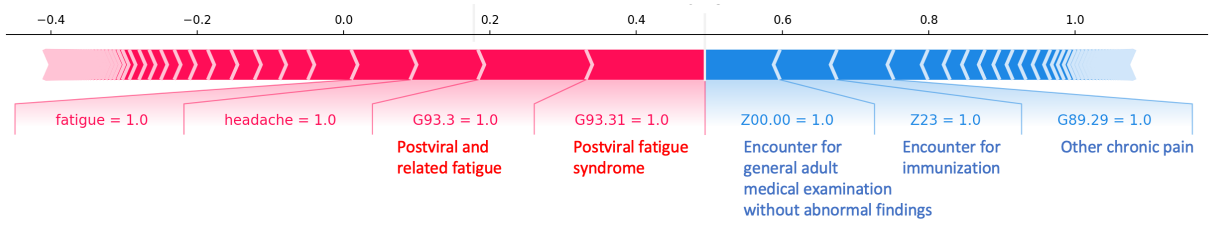

Figure 15. Example of a patient who was wrongly predicted as positive (false positive). Shown is a force plot visualizing the interactions of the positively predictive features and negatively predictive features. Fatigue, headache, G93.3 and G93.31 (red) pushed the prediction to positive, and they had a greater impact than Z00.00, Z23, and G89.29 (blue). The 1.0 value on the right of each feature indicated that the patient had the feature.

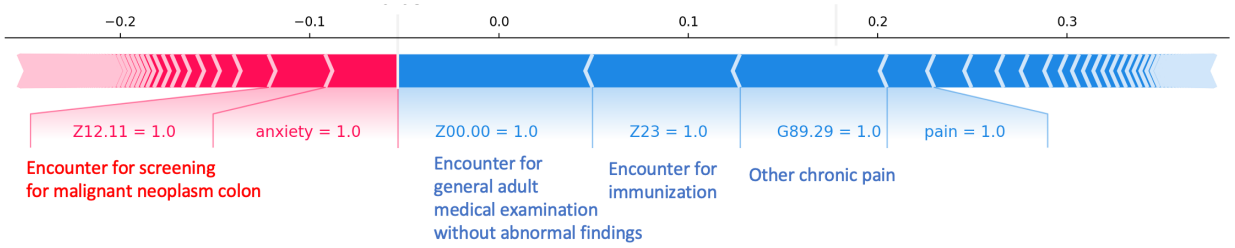

Figure 16. Example of a patient who was wrongly predicted as negative (false negative). Shown is a force plot visualizing the interactions of the positively predictive features and negatively predictive features. Z00.00, Z23, Z89.19, and pain (blue) pushed the prediction to negative, and they had a greater impact than Z12.11 and anxiety. The 1.0 value on the right of each feature indicated that the patient had the feature.

6.7 Discussion

Symptomatic differences between study groups

The results of our study can be interpreted in various ways. Our findings suggest that significant symptomatic differences exist between primary care Long COVID patients and the specialty care Long COVID patients. In practice, patients who seek care in the specialty care clinic tend to have more severe Long COVID symptoms. In contrast, the primary care Long COVID patients may recover quickly and do not require a visit to the Long COVID clinic, even if they had a record of U09.9. From this perspective, we are comparing less severe patients to more severe Long COVID patients. Our findings suggest that more severe patients were more likely to experience fatigue and cognitive issues. This might be useful for triaging Long COVID patients.

On the other hand, based on the NLP-extracted symptoms, we found fatigue was mentioned in more than 70% of the primary care group patients' clinical records, but the ICD codes only covered 20% of them. This discrepancy highlights an opportunity to inform primary care physicians about documenting Long COVID-related ICD codes more extensively.

NLP symptoms

Our analysis using NLP showed that many Long COVID-related symptoms were under-documented as ICD codes. Consequently, studies derived from ICD codes alone might be under representative. For example, in analyzing the expert derived Long COVID symptoms (ICD code based), 23% of specialty care patients were documented to have fatigue, whereas NLP indicates that the number should be 96%. Similarly, 18% of specialty care patients were documented to have shortness of breath, whereas NLP indicates that the number should be 59%.

Prediction model

Predicting Long COVID remains to be challenging. One reason could be the change of the understanding of U09.9 Long COVID overtime. Our study heavily relived on the U09.9 code to identify Long COVID patients, but the physicians' definition might change over time as they gained more understanding of Long COVID. Our dataset may not be sufficient to capture these details. The chronological aspect of visits is also crucial; patients who had frequent Long COVID related primary visits might had a higher chance of being sent to the specialty care Long COVID clinic, and our model did not account for this.

Limitations

While we have found significant symptom comparisons between the specialty care and primary care groups, we must acknowledge that the two populations differ in many ways. For example, race (White: 67% vs. 62%), insurance type (commercial: 62% vs. 39%), and age (younger vs. older) differ between the groups. These factors should be taken into account when disseminating the results. Our machine learning model only used NLP symptoms as features and future studies could considering adding the demographics information of patients which could be a significant predictor for Long COVID clinic admission.

Chapter 7. CONCLUSIONS

In this dissertation, we developed transferable NLP methods for classifying clinical note sections and addressed the data imbalance issue in coding female suicide reports using large language models. Additionally, we utilized NLP methods to understand symptom trajectories prior to OHCA and identify key symptoms that differentiate OHCA and non-OHCA patients. We also applied NLP to study Long COVID and found it captured significantly more symptoms than ICD codes, which can improve people's understanding of Long COVID.

While this thesis has made significant progress in addressing the challenges of natura language processing methods and their applications in understanding diseases, there are several limitations and areas for future research.

One of the key limitations in this thesis is that it develops natural language processing method using dataset from a variety of healthcare institutions. Our methods have shown promising results in generalization, but there may be unexplored challenges in dealing with highly specific and institution-dependent language patterns. The methods developed in this work also are also primarily focused on clinical narratives documented by healthcare workers. Other important text sources such as patient-doctor conversation could also be explored in the future.

On the application side, our focus was primarily on investigating diseases using NLP techniques applied to clinical narratives. While we can potentially made significant efforts in understanding diseases by analyzing the symptoms documented in the notes, it is important to acknowledge that diseases also have complex pathological aspects that were beyond the scope of this research. Pathological aspects encompass the underlying biological mechanisms, cellular and molecular processes, and anatomical changes associated with diseases. Future research should aim to bridge the gap between the clinical narrative analysis and the pathological features of

diseases, creating a more holistic and multidisciplinary approach to disease understanding and management.

BIBLIOGRAPHY

1 Esteva A, Robicquet A, Ramsundar B, *et al.* A guide to deep learning in healthcare. *Nat Med* 2019; **25**: 24–9.
2 Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med* 2022; **28**: 31–8.
3 Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018; **2**: 719–31.
4 Futoma J, Morris J, Lucas J. A comparison of models for predicting early hospital readmissions. *J Biomed Inform* 2015; **56**: 229–38.
5 Bajor JM, Lasko TA. Predicting Medications from Diagnostic Codes with Recurrent Neural Networks. 2022. https://openreview.net/forum?id=rJEgeXFex (accessed May 27, 2024).
6 Zhang D, Yin C, Zeng J, Yuan X, Zhang P. Combining structured and unstructured data for predictive models: a deep learning approach. *BMC Med Inform Decis Mak* 2020; **20**: 280.
7 Lybarger K, Ostendorf M, Thompson M, Yetisgen M. Extracting COVID-19 Diagnoses and Symptoms From Clinical Text: A New Annotated Corpus and Neural Event Extraction Framework. *ArXiv* 2020; : arXiv:2012.00974v1.
8 Edinger T, Demner-Fushman D, Cohen AM, Bedrick S, Hersh W. Evaluation of Clinical Text Segmentation to Facilitate Cohort Retrieval. *AMIA Annu Symp Proc* 2018; **2017**: 660–9.
9 Lei J, Tang B, Lu X, Gao K, Jiang M, Xu H. A comprehensive study of named entity recognition in Chinese clinical text. *J Am Med Inform Assoc* 2014; **21**: 808–14.
10 Nath S, Marie A, Ellershaw S, Korot E, Keane PA. New meaning for NLP: the trials and tribulations of natural language processing with GPT-3 in ophthalmology. *Br J Ophthalmol* 2022; **106**: 889–92.
11 Johnson AEW, Pollard TJ, Shen L, *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* 2016; **3**: 160035.
12 Gao Y, Dligach D, Miller T, *et al.* Hierarchical Annotation for Building A Suite of Clinical Natural Language Processing Tasks: Progress Note Understanding. *LREC Int Conf Lang Resour Eval* 2022; **2022**: 5484–93.
13 Gao Y, Caskey J, Miller T, *et al.* Tasks 1 and 3 from Progress Note Understanding Suite of Tasks: SOAP Note Tagging and Problem List Summarization. DOI:10.13026/WKS0-W041.
14 Pomares-Quimbaya A, Kreuzthaler M, Schulz S. Current approaches to identify sections within clinical narratives from electronic health records: a systematic review. *BMC Med Res Methodol* 2019; **19**: 155.
15 Beeferman D, Berger A, Lafferty J. Statistical models for text segmentation. *Machine learning* 1999; **34**: 177–210.
16 Tepper M, Capurro D, Xia F, Vanderwende L, Yetisgen-Yildiz M. Statistical Section Segmentation in Free-Text Clinical Records. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12). Istanbul, Turkey: European Language Resources Association (ELRA), 2012: 2001–8.
17 Zhou W, Yetisgen M, Afshar M, Gao Y, Savova G, Miller TA. Improving model transferability for clinical note section classification models using continued pretraining. *Journal of the American Medical Informatics Association* 2023; : ocad190.

18 Miller M, Zhang Y, Prince L, *et al.* Suicide Deaths Among Women in California Living With Handgun Owners vs Those Living With Other Adults in Handgun-Free Homes, 2004-2016. *JAMA Psychiatry* 2022; **79**: 582–8.
19 Kung TH, Cheatham M, Medenilla A, *et al.* Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health* 2023; **2**: e0000198.
20 Zhou W, Prater LC, Goldstein EV, Mooney SJ. Identifying Rare Circumstances Preceding Female Firearm Suicides: Validating A Large Language Model Approach. *JMIR Mental Health* 2023; **10**: e49359.
21 Hayashi M, Shimizu W, Albert CM. The Spectrum of Epidemiology Underlying Sudden Cardiac Death. *Circ Res* 2015; **116**: 1887–906.
22 Höglund H, Jansson J-H, Forslund A-S, Lundblad D. Prodromal symptoms and health care consumption prior to out-of-hospital cardiac arrest in patients without previously known ischaemic heart disease. *Resuscitation* 2014; **85**: 864–8.
23 Marijon E, Uy-Evanado A, Dumas F, *et al.* Warning Symptoms Are Associated With Survival From Sudden Cardiac Arrest. *Ann Intern Med* 2016; **164**: 23–9.
24 Reinier K, Dizon B, Chugh H, *et al.* Warning symptoms associated with imminent sudden cardiac arrest: a population-based case-control study with external validation. *Lancet Digit Health* 2023; **5**: e763–73.
25 Jabbari R, Risgaard B, Holst AG, *et al.* Cardiac symptoms before sudden cardiac death caused by coronary artery disease: a nationwide study among young Danish people. *Heart* 2013; **99**: 938–43.
26 Spatz ES, Gottlieb M, Wisk LE, *et al.* Three-Month Symptom Profiles Among Symptomatic Adults With Positive and Negative Severe Acute Respiratory Syndrome Coronavirus 2 Tests: A Prospective Cohort Study From the INSPIRE Group. *Clin Infect Dis* 2023; **76**: 1559–66.
27 Burkhardt HA, Mollis B, Au M, *et al.* Extracting COVID-19 Related Symptoms from EHR Data: A Comparison of Three Methods. 2021; published online March. https://familymedicine.uw.edu/stephensgroup/wp-content/uploads/sites/11/2021/07/2020-12-21-Poster-Abstract.pdf.
28 Deer RR, Rock MA, Vasilevsky N, *et al.* Characterizing Long COVID: Deep Phenotype of a Complex Condition. *EBioMedicine* 2021; **74**: 103722.
29 Sneller MC, Liang CJ, Marques AR, *et al.* A Longitudinal Study of COVID-19 Sequelae and Immunity: Baseline Findings. *Ann Intern Med* 2022; **175**: 969–79.
30 Dubin JM, Bennett NE, Halpern JA. The adverse impact of COVID-19 on men's health. *Curr Opin Urol* 2022; **32**: 146–51.
31 Zweigenbaum P, Deléger L, Lavergne T, Névéol A, Bodnari A. A Supervised Abbreviation Resolution System for Medical Text. 2013. https://www.semanticscholar.org/paper/A-Supervised-Abbreviation-Resolution-System-for-Zweigenbaum-Del%C3%A9ger/b3ba1306d0afb9f69412df1ca35ee1c49cf27a13 (accessed Feb 20, 2023).
32 Kropf S, Krücken P, Mueller W, Denecke K. Structuring Legacy Pathology Reports by openEHR Archetypes to Enable Semantic Querying. *Methods Inf Med* 2017; **56**: 230–7.
33 Rosenthal S, Barker K, Liang Z. Leveraging Medical Literature for Section Prediction in Electronic Health Records. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong, China: Association for Computational Linguistics, 2019: 4864–73.

34 Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, Minnesota: Association for Computational Linguistics, 2019: 4171–86.
35 Denny JC, Spickard A, Johnson KB, Peterson NB, Peterson JF, Miller RA. Evaluation of a method to identify and categorize section headers in clinical documents. *J Am Med Inform Assoc* 2009; **16**: 806–15.
36 Li Y, Gorman SL, Elhadad N. Section classification in clinical notes using supervised hidden markov model. Proceedings of the 1st ACM International Health Informatics Symposium. 2010; : 744–50.
37 Podder V, Lew V, Ghassemzadeh S. SOAP Notes. In: StatPearls. Treasure Island (FL): StatPearls Publishing, 2022. http://www.ncbi.nlm.nih.gov/books/NBK482263/ (accessed Jan 4, 2023).
38 Goldstein EV, Mooney SJ, Takagi-Stewart J, *et al.* Characterizing Female Firearm Suicide Circumstances: A Natural Language Processing and Machine Learning Approach. *Am J Prev Med* 2023; : S0749-3797(23)00047-8.
39 Anestis MD. Prior suicide attempts are less common in suicide decedents who died by firearms relative to those who died by other means. *Journal of Affective Disorders* 2016; **189**: 106–9.
40 Kellermann AL, Rivara FP, Somes G, *et al.* Suicide in the home in relation to gun ownership. *N Engl J Med* 1992; **327**: 467–72.
41 Centers for Disease Control and Prevention. Web-Based injury statistics query and reporting system (WISQARS). https://www.cdc.gov/injury/wisqars/index.html (accessed Feb 16, 2023).
42 Crosby AE, Ertl A, Lyons BH, Ivey-Stephenson AZ, Jack SPD. Circumstances Associated With Suicides Among Females—16 States, United States, 2005–2016. *Med Care* 2021; **59**: S92–9.
43 Choi NG, DiNitto DM, Sagna AO, Marti CN. Older women who died by suicide: suicide means, sociodemographic and psychiatric risk factors, and other precipitating circumstances. *International Psychogeriatrics* 2018; **30**: 1531–40.
44 Nichol G, Thomas E, Callaway CW, *et al.* Regional variation in out-of-hospital cardiac arrest incidence and outcome. *JAMA* 2008; **300**: 1423–31.
45 McNally B, Robb R, Mehta M, *et al.* Out-of-hospital cardiac arrest surveillance --- Cardiac Arrest Registry to Enhance Survival (CARES), United States, October 1, 2005--December 31, 2010. *MMWR Surveill Summ* 2011; **60**: 1–19.
46 Fishman GI, Chugh SS, DiMarco JP, *et al.* Sudden Cardiac Death Prediction and Prevention Report From a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop. *Circulation* 2010; **122**: 2335–48.
47 Nishiyama C, Iwami T, Kawamura T, *et al.* Prodromal symptoms of out-of-hospital cardiac arrests: a report from a large-scale population-based cohort study. *Resuscitation* 2013; **84**: 558–63.
48 Kreimeyer K, Foster M, Pandey A, *et al.* Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform* 2017; **73**: 14–29.

49 Wang L, Foer D, MacPhaul E, Lo Y-C, Bates DW, Zhou L. PASCLex: A comprehensive post-acute sequelae of COVID-19 (PASC) symptom lexicon derived from electronic health record clinical notes. *J Biomed Inform* 2022; **125**: 103951.
50 Pfaff ER, Girvin AT, Bennett TD, *et al.* Identifying who has long COVID in the USA: a machine learning approach using N3C data. *The Lancet Digital Health* 2022; **4**: e532–41.
51 Thaweethai T, Jolley SE, Karlson EW, *et al.* Development of a Definition of Postacute Sequelae of SARS-CoV-2 Infection. *JAMA* 2023; **329**: 1934–46.
52 Wright A, Sittig DF, McGowan J, Ash JS, Weed LL. Bringing science to medicine: an interview with Larry Weed, inventor of the problem-oriented medical record. *J Am Med Inform Assoc* 2014; **21**: 964–8.
53 Gururangan S, Marasović A, Swayamdipta S, *et al.* Don't Stop Pretraining: Adapt Language Models to Domains and Tasks. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, 2020: 8342–60.
54 Iv WFS, Bethard S, Finan S, *et al.* Temporal Annotation in the Clinical Domain. *Transactions of the Association for Computational Linguistics* 2014; **2**: 143–54.
55 Mowery D, Wiebe J, Visweswaran S, Harkema H, Chapman WW. Building an automated SOAP classifier for emergency department reports. *J Biomed Inform* 2012; **45**: 71–81.
56 Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics* 2008; **77**: 291–304.
57 Lee J, Yoon W, Kim S, *et al.* BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* 2020; **36**: 1234–40.
58 Alsentzer E, Murphy J, Boag W, *et al.* Publicly Available Clinical BERT Embeddings. In: Proceedings of the 2nd Clinical Natural Language Processing Workshop. Minneapolis, Minnesota, USA: Association for Computational Linguistics, 2019: 72–8.
59 Wolf T, Debut L, Sanh V, *et al.* HuggingFace's Transformers: State-of-the-art Natural Language Processing. 2020; published online July 13. DOI:10.48550/arXiv.1910.03771.
60 Wu S, Miller T, Masanz J, *et al.* Negation's Not Solved: Generalizability Versus Optimizability in Clinical Natural Language Processing. *PLOS ONE* 2014; **9**: e112774.
61 Xu FF, Alon U, Neubig G, Hellendoorn VJ. A systematic evaluation of large language models of code. In: Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming. New York, NY, USA: Association for Computing Machinery, 2022: 1–10.
62 Chung HW, Hou L, Longpre S, *et al.* Scaling Instruction-Finetuned Language Models. 2022; published online Dec 6. DOI:10.48550/arXiv.2210.11416.
63 A New Open Source Flan 20B with UL2. Yi Tay. https://www.yitay.net/blog/flan-ul2-20b (accessed April 17, 2023).
64 Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press, 2016.
65 Myat A, Song K-J, Rea T. Out-of-hospital cardiac arrest: current concepts. *Lancet* 2018; **391**: 970–9.
66 Zhou S, Lybarger K, Yetisgen M, Ostendorf M. Generalizing through Forgetting - Domain Generalization for Symptom Event Extraction in Clinical Notes. *AMIA Jt Summits Transl Sci Proc* 2023; **2023**: 622–31.

67 Prado MG, Kessler LG, Au MA, *et al.* Symptoms and signs of lung cancer prior to diagnosis: case-control study using electronic health records from ambulatory care within a large US-based tertiary care centre. *BMJ Open* 2023; **13**: e068832.
68 Bobak L, Dorney I, Kovacevich A, Barnett B, Kaelber DC. Preexisting Psychiatric Conditions as Risk Factors for Diagnosed Long COVID-19 Syndrome Within Aggregated Electronic Health Record Data. *Psychosom Med* 2024; **86**: 132–6.
69 Glasheen WP, Cordier T, Gumpina R, Haugh G, Davis J, Renda A. Charlson Comorbidity Index: ICD-9 Update and ICD-10 Translation. *Am Health Drug Benefits* 2019; **12**: 188–97.
70 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 4768–77.

APPENDICES

Table 1. Symptom prevalence by symptom and time prior to OHCA.

Symptom	Days Prior to OHCA	Total Cohort N	Prevalence Symptom in OHCA (%)	Prevalence Symptom in Control (%)
Bleeding	30	18677	20.1	9.9
Chest pain	30	18677	28.3	15.9
Diaphoresis	30	18677	4.7	2.1
Headache	30	18677	21.4	16.1
Viral respiratory symptoms	30	18677	59.1	41.2
Peripheral edema	30	18677	23.4	8.8
Light headedness	30	18677	21.6	11.9
Nausea	30	18677	25	13.9
Palpitations	30	18677	9.9	5.5
Seizure	30	18677	7.4	2.3
Shortness of breath	30	18677	36.6	16.6
Insomnia	30	18677	20.9	12.3
Syncope	30	18677	9.1	4
Bleeding	60	18677	19	9
Chest pain	60	18677	27	14.7
Diaphoresis	60	18677	4.5	2
Headache	60	18677	20	15.1
Viral respiratory symptoms	60	18677	56.2	38.5
Peripheral edema	60	18677	22.2	8.2
Light headedness	60	18677	20.6	11.1
Nausea	60	18677	23	12.9
Palpitations	60	18677	9	5
Seizure	60	18677	6.8	2.2
Shortness of breath	60	18677	34.5	15.5
Insomnia	60	18677	19.3	11.6
Syncope	60	18677	8.4	3.7
Bleeding	90	18677	17.9	8.5
Chest pain	90	18677	25.9	13.8
Diaphoresis	90	18677	4.2	1.9
Headache	90	18677	18.9	14.3
Viral respiratory symptoms	90	18677	53.2	36.4
Peripheral edema	90	18677	21.2	7.7

Light headedness	90	18677	19.9	10.3
Nausea	90	18677	22	12
Palpitations	90	18677	8.8	4.8
Seizure	90	18677	6.6	2
Shortness of breath	90	18677	33	14.5
Insomnia	90	18677	18.3	10.9
Syncope	90	18677	7.7	3.5
Bleeding	120	18677	17.1	8.1
Chest pain	120	18677	24.5	13.1
Diaphoresis	120	18677	4.1	1.8
Headache	120	18677	18.1	13.5
Viral respiratory symptoms	120	18677	51.5	34.5
Peripheral edema	120	18677	20.4	7.1
Light headedness	120	18677	18.4	9.9
Nausea	120	18677	20.9	11.3
Palpitations	120	18677	8.5	4.6
Seizure	120	18677	6.3	1.9
Shortness of breath	120	18677	31.3	13.7
Insomnia	120	18677	18	10.3
Syncope	120	18677	7.4	3.2
Bleeding	150	18677	16.5	7.6
Chest pain	150	18677	23.5	12.5
Diaphoresis	150	18677	4	1.7
Headache	150	18677	17.3	12.8
Viral respiratory symptoms	150	18677	50.1	32.9
Peripheral edema	150	18677	19.6	6.7
Light headedness	150	18677	17.8	9.3
Nausea	150	18677	20.3	10.6
Palpitations	150	18677	8.1	4.4
Seizure	150	18677	6.3	1.8
Shortness of breath	150	18677	29.8	12.9
Insomnia	150	18677	17.4	9.8
Syncope	150	18677	7.3	3
Bleeding	180	18677	15.7	7.2
Chest pain	180	18677	22.5	11.9
Diaphoresis	180	18677	4	1.6
Headache	180	18677	17	12.2
Viral respiratory symptoms	180	18677	48.9	31.3

Peripheral edema	180	18677	18.9	6.2
Light headedness	180	18677	17.4	8.8
Nausea	180	18677	19.4	9.9
Palpitations	180	18677	7.9	4.1
Seizure	180	18677	6	1.7
Shortness of breath	180	18677	28.8	12.2
Insomnia	180	18677	16.5	9.4
Syncope	180	18677	6.7	2.9
Bleeding	210	18677	14.5	6.8
Chest pain	210	18677	21.3	11.2
Diaphoresis	210	18677	3.8	1.6
Headache	210	18677	16.1	11.6
Viral respiratory symptoms	210	18677	46.8	29.9
Peripheral edema	210	18677	18.1	5.8
Light headedness	210	18677	16.2	8.3
Nausea	210	18677	17.9	9.3
Palpitations	210	18677	7.5	3.9
Seizure	210	18677	5.6	1.6
Shortness of breath	210	18677	27.7	11.5
Insomnia	210	18677	15.8	8.9
Syncope	210	18677	6.5	2.7
Bleeding	240	18677	14	6.3
Chest pain	240	18677	20.4	10.7
Diaphoresis	240	18677	3.6	1.5
Headache	240	18677	15.4	11
Viral respiratory symptoms	240	18677	45.2	28.6
Peripheral edema	240	18677	17.1	5.4
Light headedness	240	18677	15.2	7.8
Nausea	240	18677	16.7	8.7
Palpitations	240	18677	7.1	3.7
Seizure	240	18677	5.1	1.5
Shortness of breath	240	18677	26.3	11
Insomnia	240	18677	15.1	8.4
Syncope	240	18677	6.3	2.6
Bleeding	270	18677	13.5	5.9
Chest pain	270	18677	19.1	10.3
Diaphoresis	270	18677	3.3	1.4
Headache	270	18677	14.7	10.5

Viral respiratory symptoms	270	18677	43.6	27.3
Peripheral edema	270	18677	16.3	5
Light headedness	270	18677	14.7	7.5
Nausea	270	18677	15.8	8.1
Palpitations	270	18677	6.9	3.6
Seizure	270	18677	4.8	1.5
Shortness of breath	270	18677	25.2	10.4
Insomnia	270	18677	14.5	8
Syncope	270	18677	6	2.5
Bleeding	300	18677	12.5	5.6
Chest pain	300	18677	17.9	9.7
Diaphoresis	300	18677	3.1	1.3
Headache	300	18677	14	10
Viral respiratory symptoms	300	18677	41.8	26.2
Peripheral edema	300	18677	15.1	4.7
Light headedness	300	18677	13.7	7.1
Nausea	300	18677	14.7	7.7
Palpitations	300	18677	6.7	3.4
Seizure	300	18677	4.4	1.4
Shortness of breath	300	18677	24	9.9
Insomnia	300	18677	13.8	7.5
Syncope	300	18677	5.6	2.4
Bleeding	330	18677	11.7	5.2
Chest pain	330	18677	17.5	9.1
Diaphoresis	330	18677	2.8	1.2
Headache	330	18677	13.5	9.4
Viral respiratory symptoms	330	18677	40.2	25
Peripheral edema	330	18677	14.5	4.4
Light headedness	330	18677	13.3	6.9
Nausea	330	18677	14.3	7.2
Palpitations	330	18677	6.1	3.2
Seizure	330	18677	4.2	1.3
Shortness of breath	330	18677	23.1	9.3
Insomnia	330	18677	13.1	7.1
Syncope	330	18677	5.3	2.3
Bleeding	360	18677	11.2	4.8
Chest pain	360	18677	16.3	8.6
Diaphoresis	360	18677	2.6	1.1

Headache	360	18677	12.6	8.8
Viral respiratory symptoms	360	18677	38.6	23.8
Peripheral edema	360	18677	13.3	4.1
Light headedness	360	18677	12.3	6.5
Nausea	360	18677	13.4	6.7
Palpitations	360	18677	5.9	3
Seizure	360	18677	3.9	1.2
Shortness of breath	360	18677	22	8.7
Insomnia	360	18677	12.1	6.7
Syncope	360	18677	5	2.1

Table 2. Symptom prevalence by symptom, subgroup and time prior to OHCA.

Subgroup	**Symptom**	**Days Prior to OHCA**	**Subgroup N**	**Prevalence Symptom in OHCA (%)**	**Prevalence of Symptoms in Controls (%)**
Male	Bleeding	30	11966	18.7	8.3
Male	Chest pain	30	11966	27.4	15.2
Male	Diaphoresis	30	11966	3.8	2
Male	Headache	30	11966	18.1	13.1
Male	Viral respiratory symptoms	30	11966	56.1	39
Male	Peripheral edema	30	11966	21.4	8.5
Male	Light headedness	30	11966	22.3	10.8
Male	Nausea	30	11966	21.2	11.3
Male	Palpitations	30	11966	8.2	4.5
Male	Seizure	30	11966	7.8	2.3
Male	Shortness of breath	30	11966	35.7	16.1
Male	Insomnia	30	11966	18.3	10.9
Male	Syncope	30	11966	8.9	3.8
Male	Bleeding	60	11966	17.6	7.6
Male	Chest pain	60	11966	26.1	14
Male	Diaphoresis	60	11966	3.7	1.8
Male	Headache	60	11966	16.6	12.4
Male	Viral respiratory symptoms	60	11966	53	36.4
Male	Peripheral edema	60	11966	20.3	8

Male	Light headedness	60	11966	21.4	10
Male	Nausea	60	11966	19.2	10.5
Male	Palpitations	60	11966	7.5	4.2
Male	Seizure	60	11966	7.2	2.1
Male	Shortness of breath	60	11966	33	15
Male	Insomnia	60	11966	16.8	10.1
Male	Syncope	60	11966	8.2	3.6
Male	Bleeding	90	11966	16.8	7.1
Male	Chest pain	90	11966	25.2	13.1
Male	Diaphoresis	90	11966	3.4	1.8
Male	Headache	90	11966	15.4	11.8
Male	Viral respiratory symptoms	90	11966	50.4	34.3
Male	Peripheral edema	90	11966	19.6	7.5
Male	Light headedness	90	11966	20.6	9.2
Male	Nausea	90	11966	18.5	9.8
Male	Palpitations	90	11966	7.2	4
Male	Seizure	90	11966	7	2
Male	Shortness of breath	90	11966	31.5	14
Male	Insomnia	90	11966	16.1	9.5
Male	Syncope	90	11966	7.6	3.3
Male	Bleeding	120	11966	16.1	6.8
Male	Chest pain	120	11966	24	12.4
Male	Diaphoresis	120	11966	3.3	1.7
Male	Headache	120	11966	14.7	11
Male	Viral respiratory symptoms	120	11966	48.7	32.3
Male	Peripheral edema	120	11966	18.5	6.9
Male	Light headedness	120	11966	19.3	8.8
Male	Nausea	120	11966	17.3	9.2
Male	Palpitations	120	11966	7	3.8
Male	Seizure	120	11966	6.8	1.8
Male	Shortness of breath	120	11966	29.6	13.2
Male	Insomnia	120	11966	15.7	9
Male	Syncope	120	11966	7.2	3.1
Male	Bleeding	150	11966	15.7	6.4
Male	Chest pain	150	11966	23.3	11.9
Male	Diaphoresis	150	11966	3.3	1.6

Male	Headache	150	11966	14	10.3
Male	Viral respiratory symptoms	150	11966	46.8	30.7
Male	Peripheral edema	150	11966	17.7	6.5
Male	Light headedness	150	11966	18.6	8.2
Male	Nausea	150	11966	16.5	8.5
Male	Palpitations	150	11966	6.7	3.6
Male	Seizure	150	11966	6.8	1.7
Male	Shortness of breath	150	11966	28.2	12.4
Male	Insomnia	150	11966	15	8.5
Male	Syncope	150	11966	7.1	2.9
Male	Bleeding	180	11966	14.9	6
Male	Chest pain	180	11966	22.4	11.3
Male	Diaphoresis	180	11966	3.3	1.5
Male	Headache	180	11966	13.6	9.8
Male	Viral respiratory symptoms	180	11966	45.6	29.1
Male	Peripheral edema	180	11966	16.9	6
Male	Light headedness	180	11966	18.1	7.7
Male	Nausea	180	11966	16	7.9
Male	Palpitations	180	11966	6.6	3.4
Male	Seizure	180	11966	6.5	1.7
Male	Shortness of breath	180	11966	27.1	11.8
Male	Insomnia	180	11966	14.3	8.2
Male	Syncope	180	11966	6.6	2.8
Male	Bleeding	210	11966	13.7	5.6
Male	Chest pain	210	11966	21.3	10.6
Male	Diaphoresis	210	11966	3.3	1.4
Male	Headache	210	11966	13.2	9.4
Male	Viral respiratory symptoms	210	11966	43.8	27.8
Male	Peripheral edema	210	11966	16.2	5.6
Male	Light headedness	210	11966	16.8	7.4
Male	Nausea	210	11966	15	7.3
Male	Palpitations	210	11966	6.5	3.3
Male	Seizure	210	11966	6.2	1.6
Male	Shortness of breath	210	11966	26.1	11
Male	Insomnia	210	11966	14	7.7
Male	Syncope	210	11966	6.4	2.6

Male	Bleeding	240	11966	13.3	5.2
Male	Chest pain	240	11966	20.3	10.2
Male	Diaphoresis	240	11966	3.1	1.4
Male	Headache	240	11966	12.6	9
Male	Viral respiratory symptoms	240	11966	42.4	26.6
Male	Peripheral edema	240	11966	15.7	5.2
Male	Light headedness	240	11966	15.8	7.1
Male	Nausea	240	11966	14	6.9
Male	Palpitations	240	11966	6.2	3.1
Male	Seizure	240	11966	5.8	1.5
Male	Shortness of breath	240	11966	24.7	10.6
Male	Insomnia	240	11966	13.2	7.3
Male	Syncope	240	11966	6.1	2.5
Male	Bleeding	270	11966	12.7	4.8
Male	Chest pain	270	11966	19	9.6
Male	Diaphoresis	270	11966	2.8	1.3
Male	Headache	270	11966	12	8.6
Male	Viral respiratory symptoms	270	11966	40.8	25.3
Male	Peripheral edema	270	11966	15	4.7
Male	Light headedness	270	11966	15.1	6.7
Male	Nausea	270	11966	13.5	6.3
Male	Palpitations	270	11966	6.1	3
Male	Seizure	270	11966	5.3	1.5
Male	Shortness of breath	270	11966	23.4	10
Male	Insomnia	270	11966	12.6	6.9
Male	Syncope	270	11966	5.8	2.4
Male	Bleeding	300	11966	11.5	4.6
Male	Chest pain	300	11966	17.8	9.1
Male	Diaphoresis	300	11966	2.8	1.2
Male	Headache	300	11966	11.7	8.1
Male	Viral respiratory symptoms	300	11966	38.9	24.3
Male	Peripheral edema	300	11966	14.2	4.4
Male	Light headedness	300	11966	14.1	6.4
Male	Nausea	300	11966	12.8	6
Male	Palpitations	300	11966	5.8	2.8
Male	Seizure	300	11966	4.9	1.4

Male	Shortness of breath	300	11966	22.4	9.5
Male	Insomnia	300	11966	11.9	6.5
Male	Syncope	300	11966	5.5	2.3
Male	Bleeding	330	11966	10.6	4.3
Male	Chest pain	330	11966	17.3	8.6
Male	Diaphoresis	330	11966	2.4	1.2
Male	Headache	330	11966	11.5	7.6
Male	Viral respiratory symptoms	330	11966	37.6	23.1
Male	Peripheral edema	330	11966	14	4.2
Male	Light headedness	330	11966	13.5	6.2
Male	Nausea	330	11966	12.5	5.6
Male	Palpitations	330	11966	5.2	2.6
Male	Seizure	330	11966	4.7	1.3
Male	Shortness of breath	330	11966	21.7	8.9
Male	Insomnia	330	11966	11.4	6.1
Male	Syncope	330	11966	5.2	2.1
Male	Bleeding	360	11966	10	3.9
Male	Chest pain	360	11966	15.8	8.1
Male	Diaphoresis	360	11966	2.3	1.1
Male	Headache	360	11966	10.8	7.2
Male	Viral respiratory symptoms	360	11966	36.2	21.9
Male	Peripheral edema	360	11966	12.6	3.8
Male	Light headedness	360	11966	13	5.9
Male	Nausea	360	11966	11.8	5.2
Male	Palpitations	360	11966	5	2.4
Male	Seizure	360	11966	4.3	1.2
Male	Shortness of breath	360	11966	20.6	8.4
Male	Insomnia	360	11966	10.4	5.8
Male	Syncope	360	11966	5	2
Female	Bleeding	30	6711	22.6	12.7
Female	Chest pain	30	6711	30	17.2
Female	Diaphoresis	30	6711	6.4	2.4
Female	Headache	30	6711	27.6	21.4
Female	Viral respiratory symptoms	30	6711	64.8	45.1
Female	Peripheral edema	30	6711	27.2	9.4
Female	Light headedness	30	6711	20.4	13.9

Female	Nausea	30	6711	32.2	18.5
Female	Palpitations	30	6711	13	7.2
Female	Seizure	30	6711	6.6	2.4
Female	Shortness of breath	30	6711	38.4	17.6
Female	Insomnia	30	6711	25.8	15
Female	Syncope	30	6711	9.4	4.3
Female	Bleeding	60	6711	21.6	11.6
Female	Chest pain	60	6711	28.8	16.1
Female	Diaphoresis	60	6711	6	2.3
Female	Headache	60	6711	26.4	20.1
Female	Viral respiratory symptoms	60	6711	62	42.3
Female	Peripheral edema	60	6711	25.8	8.6
Female	Light headedness	60	6711	19.2	13
Female	Nausea	60	6711	30.2	17.3
Female	Palpitations	60	6711	12	6.6
Female	Seizure	60	6711	6	2.2
Female	Shortness of breath	60	6711	37.2	16.4
Female	Insomnia	60	6711	24	14.1
Female	Syncope	60	6711	8.8	3.8
Female	Bleeding	90	6711	20	10.9
Female	Chest pain	90	6711	27.2	15
Female	Diaphoresis	90	6711	5.8	2.2
Female	Headache	90	6711	25.4	18.8
Female	Viral respiratory symptoms	90	6711	58.4	40.1
Female	Peripheral edema	90	6711	24.2	8
Female	Light headedness	90	6711	18.6	12.3
Female	Nausea	90	6711	28.4	16
Female	Palpitations	90	6711	11.6	6.3
Female	Seizure	90	6711	5.8	2
Female	Shortness of breath	90	6711	35.8	15.3
Female	Insomnia	90	6711	22.6	13.4
Female	Syncope	90	6711	8	3.6
Female	Bleeding	120	6711	19	10.4
Female	Chest pain	120	6711	25.6	14.2
Female	Diaphoresis	120	6711	5.6	2.1
Female	Headache	120	6711	24.4	18

Female	Viral respiratory symptoms	120	6711	56.8	38.4
Female	Peripheral edema	120	6711	24	7.5
Female	Light headedness	120	6711	16.8	11.8
Female	Nausea	120	6711	27.8	15.1
Female	Palpitations	120	6711	11.4	6.1
Female	Seizure	120	6711	5.4	1.9
Female	Shortness of breath	120	6711	34.6	14.5
Female	Insomnia	120	6711	22.4	12.8
Female	Syncope	120	6711	7.8	3.5
Female	Bleeding	150	6711	18.2	9.8
Female	Chest pain	150	6711	23.8	13.5
Female	Diaphoresis	150	6711	5.4	2
Female	Headache	150	6711	23.6	17.1
Female	Viral respiratory symptoms	150	6711	56.4	36.7
Female	Peripheral edema	150	6711	23.2	7
Female	Light headedness	150	6711	16.2	11.1
Female	Nausea	150	6711	27.4	14.3
Female	Palpitations	150	6711	10.8	5.8
Female	Seizure	150	6711	5.2	1.9
Female	Shortness of breath	150	6711	32.8	13.7
Female	Insomnia	150	6711	21.8	12.2
Female	Syncope	150	6711	7.6	3.3
Female	Bleeding	180	6711	17.2	9.2
Female	Chest pain	180	6711	22.8	12.9
Female	Diaphoresis	180	6711	5.2	1.9
Female	Headache	180	6711	23.4	16.3
Female	Viral respiratory symptoms	180	6711	55.2	35.2
Female	Peripheral edema	180	6711	22.6	6.5
Female	Light headedness	180	6711	16.2	10.6
Female	Nausea	180	6711	25.8	13.5
Female	Palpitations	180	6711	10.2	5.3
Female	Seizure	180	6711	5	1.7
Female	Shortness of breath	180	6711	32	13
Female	Insomnia	180	6711	20.8	11.5
Female	Syncope	180	6711	7	3.1
Female	Bleeding	210	6711	16	8.9

Female	Chest pain	210	6711	21.4	12.3
Female	Diaphoresis	210	6711	4.6	1.8
Female	Headache	210	6711	21.6	15.5
Female	Viral respiratory symptoms	210	6711	52.4	33.5
Female	Peripheral edema	210	6711	21.6	6.1
Female	Light headedness	210	6711	15	9.8
Female	Nausea	210	6711	23.2	12.8
Female	Palpitations	210	6711	9.4	5.1
Female	Seizure	210	6711	4.6	1.6
Female	Shortness of breath	210	6711	30.8	12.3
Female	Insomnia	210	6711	19.4	11
Female	Syncope	210	6711	6.8	3
Female	Bleeding	240	6711	15.4	8.2
Female	Chest pain	240	6711	20.4	11.7
Female	Diaphoresis	240	6711	4.6	1.6
Female	Headache	240	6711	20.8	14.7
Female	Viral respiratory symptoms	240	6711	50.4	32.1
Female	Peripheral edema	240	6711	19.8	5.7
Female	Light headedness	240	6711	14.2	9.2
Female	Nausea	240	6711	22	12
Female	Palpitations	240	6711	8.8	4.9
Female	Seizure	240	6711	4	1.5
Female	Shortness of breath	240	6711	29.2	11.7
Female	Insomnia	240	6711	18.6	10.4
Female	Syncope	240	6711	6.8	2.9
Female	Bleeding	270	6711	15	7.8
Female	Chest pain	270	6711	19.4	11.4
Female	Diaphoresis	270	6711	4.4	1.5
Female	Headache	270	6711	19.6	14
Female	Viral respiratory symptoms	270	6711	49	30.9
Female	Peripheral edema	270	6711	18.8	5.5
Female	Light headedness	270	6711	13.8	8.9
Female	Nausea	270	6711	20	11.3
Female	Palpitations	270	6711	8.6	4.7
Female	Seizure	270	6711	3.8	1.5
Female	Shortness of breath	270	6711	28.6	11.1

Female	Insomnia	270	6711	18.2	9.9
Female	Syncope	270	6711	6.4	2.7
Female	Bleeding	300	6711	14.4	7.5
Female	Chest pain	300	6711	18.2	10.7
Female	Diaphoresis	300	6711	3.8	1.4
Female	Headache	300	6711	18.2	13.3
Female	Viral respiratory symptoms	300	6711	47.2	29.5
Female	Peripheral edema	300	6711	16.8	5.2
Female	Light headedness	300	6711	13	8.4
Female	Nausea	300	6711	18.4	10.6
Female	Palpitations	300	6711	8.4	4.6
Female	Seizure	300	6711	3.4	1.4
Female	Shortness of breath	300	6711	27	10.5
Female	Insomnia	300	6711	17.4	9.3
Female	Syncope	300	6711	5.8	2.6
Female	Bleeding	330	6711	13.8	7
Female	Chest pain	330	6711	18	10
Female	Diaphoresis	330	6711	3.6	1.3
Female	Headache	330	6711	17.2	12.6
Female	Viral respiratory symptoms	330	6711	45.2	28.3
Female	Peripheral edema	330	6711	15.4	4.8
Female	Light headedness	330	6711	12.8	8.1
Female	Nausea	330	6711	17.8	10.1
Female	Palpitations	330	6711	7.8	4.3
Female	Seizure	330	6711	3.2	1.4
Female	Shortness of breath	330	6711	25.6	9.9
Female	Insomnia	330	6711	16.2	8.8
Female	Syncope	330	6711	5.4	2.5
Female	Bleeding	360	6711	13.4	6.5
Female	Chest pain	360	6711	17.2	9.5
Female	Diaphoresis	360	6711	3.2	1.2
Female	Headache	360	6711	16.2	11.8
Female	Viral respiratory symptoms	360	6711	43	27
Female	Peripheral edema	360	6711	14.8	4.6
Female	Light headedness	360	6711	11	7.6
Female	Nausea	360	6711	16.4	9.3

Female	Palpitations	360	6711	7.6	4
Female	Seizure	360	6711	3.2	1.3
Female	Shortness of breath	360	6711	24.6	9.2
Female	Insomnia	360	6711	15.2	8.3
Female	Syncope	360	6711	5	2.4
White	Bleeding	30	12756	19.9	11.2
White	Chest pain	30	12756	27.9	17.2
White	Diaphoresis	30	12756	4.8	2.2
White	Headache	30	12756	21.4	17.5
White	Viral respiratory symptoms	30	12756	60.4	44.6
White	Peripheral edema	30	12756	24.5	10
White	Light headedness	30	12756	22.2	13.1
White	Nausea	30	12756	25.8	15.3
White	Palpitations	30	12756	10.3	6.1
White	Seizure	30	12756	7	2.7
White	Shortness of breath	30	12756	37.5	18.5
White	Insomnia	30	12756	21.6	13.8
White	Syncope	30	12756	8.9	4.6
White	Bleeding	60	12756	18.9	10.2
White	Chest pain	60	12756	26.8	16.1
White	Diaphoresis	60	12756	4.7	2.1
White	Headache	60	12756	20	16.5
White	Viral respiratory symptoms	60	12756	57.2	41.9
White	Peripheral edema	60	12756	22.9	9.3
White	Light headedness	60	12756	21	12.2
White	Nausea	60	12756	23.6	14.4
White	Palpitations	60	12756	9.3	5.7
White	Seizure	60	12756	6.5	2.5
White	Shortness of breath	60	12756	35.2	17.3
White	Insomnia	60	12756	19.9	13
White	Syncope	60	12756	8	4.3
White	Bleeding	90	12756	18	9.6
White	Chest pain	90	12756	25.7	15.1
White	Diaphoresis	90	12756	4.6	2.1
White	Headache	90	12756	18.8	15.6
White	Viral respiratory symptoms	90	12756	53.6	39.7

White	Peripheral edema	90	12756	21.6	8.7
White	Light headedness	90	12756	20.2	11.4
White	Nausea	90	12756	22.7	13.4
White	Palpitations	90	12756	9.1	5.4
White	Seizure	90	12756	6.4	2.3
White	Shortness of breath	90	12756	33.5	16.2
White	Insomnia	90	12756	19.1	12.3
White	Syncope	90	12756	7.2	4
White	Bleeding	120	12756	17	9.2
White	Chest pain	120	12756	24.6	14.3
White	Diaphoresis	120	12756	4.6	2
White	Headache	120	12756	18.2	14.8
White	Viral respiratory symptoms	120	12756	51.6	37.7
White	Peripheral edema	120	12756	21.2	8.1
White	Light headedness	120	12756	18.5	10.9
White	Nausea	120	12756	21.7	12.6
White	Palpitations	120	12756	8.9	5.2
White	Seizure	120	12756	6.1	2.2
White	Shortness of breath	120	12756	31.3	15.4
White	Insomnia	120	12756	18.8	11.6
White	Syncope	120	12756	6.8	3.7
White	Bleeding	150	12756	16.5	8.7
White	Chest pain	150	12756	23.5	13.7
White	Diaphoresis	150	12756	4.5	1.9
White	Headache	150	12756	17.2	14
White	Viral respiratory symptoms	150	12756	49.9	36
White	Peripheral edema	150	12756	20.1	7.6
White	Light headedness	150	12756	18	10.2
White	Nausea	150	12756	21	11.9
White	Palpitations	150	12756	8.5	5
White	Seizure	150	12756	6.1	2.1
White	Shortness of breath	150	12756	29.3	14.5
White	Insomnia	150	12756	18.3	11.1
White	Syncope	150	12756	6.8	3.5
White	Bleeding	180	12756	15.8	8.2
White	Chest pain	180	12756	22.3	13.1

White	Diaphoresis	180	12756	4.3	1.8
White	Headache	180	12756	17.1	13.3
White	Viral respiratory symptoms	180	12756	49	34.3
White	Peripheral edema	180	12756	19.3	7.1
White	Light headedness	180	12756	17.7	9.7
White	Nausea	180	12756	20.4	11.1
White	Palpitations	180	12756	8.2	4.7
White	Seizure	180	12756	5.9	2
White	Shortness of breath	180	12756	28.4	13.8
White	Insomnia	180	12756	17.2	10.6
White	Syncope	180	12756	6.1	3.4
White	Bleeding	210	12756	14.7	7.7
White	Chest pain	210	12756	21.1	12.4
White	Diaphoresis	210	12756	4.1	1.7
White	Headache	210	12756	16	12.7
White	Viral respiratory symptoms	210	12756	46.5	32.8
White	Peripheral edema	210	12756	18.5	6.6
White	Light headedness	210	12756	16.3	9.1
White	Nausea	210	12756	18.8	10.4
White	Palpitations	210	12756	7.8	4.5
White	Seizure	210	12756	5.3	1.9
White	Shortness of breath	210	12756	27.2	13
White	Insomnia	210	12756	16.4	10
White	Syncope	210	12756	6	3.2
White	Bleeding	240	12756	14.1	7.1
White	Chest pain	240	12756	20	11.8
White	Diaphoresis	240	12756	4	1.6
White	Headache	240	12756	15.4	12
White	Viral respiratory symptoms	240	12756	44.9	31.3
White	Peripheral edema	240	12756	17.4	6.2
White	Light headedness	240	12756	15.3	8.6
White	Nausea	240	12756	17.6	9.7
White	Palpitations	240	12756	7.6	4.3
White	Seizure	240	12756	4.9	1.8
White	Shortness of breath	240	12756	26.2	12.4
White	Insomnia	240	12756	15.8	9.4

White	Syncope	240	12756	5.9	3.1
White	Bleeding	270	12756	13.5	6.7
White	Chest pain	270	12756	19.3	11.4
White	Diaphoresis	270	12756	3.8	1.5
White	Headache	270	12756	14.6	11.4
White	Viral respiratory symptoms	270	12756	43.5	29.9
White	Peripheral edema	270	12756	16.7	5.7
White	Light headedness	270	12756	15	8.2
White	Nausea	270	12756	16.8	9.1
White	Palpitations	270	12756	7.6	4.1
White	Seizure	270	12756	4.5	1.7
White	Shortness of breath	270	12756	25.5	11.8
White	Insomnia	270	12756	15.2	8.9
White	Syncope	270	12756	5.5	2.9
White	Bleeding	300	12756	12.6	6.4
White	Chest pain	300	12756	18.4	10.8
White	Diaphoresis	300	12756	3.6	1.4
White	Headache	300	12756	14.2	10.9
White	Viral respiratory symptoms	300	12756	41.7	28.7
White	Peripheral edema	300	12756	15.4	5.3
White	Light headedness	300	12756	13.8	7.8
White	Nausea	300	12756	15.9	8.6
White	Palpitations	300	12756	7.3	3.9
White	Seizure	300	12756	3.9	1.6
White	Shortness of breath	300	12756	24.1	11.2
White	Insomnia	300	12756	14.2	8.5
White	Syncope	300	12756	5.1	2.8
White	Bleeding	330	12756	11.7	5.9
White	Chest pain	330	12756	17.9	10.1
White	Diaphoresis	330	12756	3.3	1.3
White	Headache	330	12756	13.9	10.3
White	Viral respiratory symptoms	330	12756	40.4	27.4
White	Peripheral edema	330	12756	14.9	5
White	Light headedness	330	12756	13.5	7.6
White	Nausea	330	12756	15.3	8.1
White	Palpitations	330	12756	6.6	3.6

White	Seizure	330	12756	3.7	1.6
White	Shortness of breath	330	12756	23	10.5
White	Insomnia	330	12756	13.4	8
White	Syncope	330	12756	4.8	2.7
White	Bleeding	360	12756	11.2	5.5
White	Chest pain	360	12756	16.6	9.6
White	Diaphoresis	360	12756	3	1.2
White	Headache	360	12756	12.8	9.7
White	Viral respiratory symptoms	360	12756	38.2	26.1
White	Peripheral edema	360	12756	13.6	4.7
White	Light headedness	360	12756	12.6	7.1
White	Nausea	360	12756	14.1	7.5
White	Palpitations	360	12756	6.4	3.4
White	Seizure	360	12756	3.4	1.5
White	Shortness of breath	360	12756	22.1	9.8
White	Insomnia	360	12756	12.6	7.6
White	Syncope	360	12756	4.5	2.5
Non-White	Bleeding	30	3155	21.6	10.2
Non-White	Chest pain	30	3155	31.3	17.9
Non-White	Diaphoresis	30	3155	4.7	2.8
Non-White	Headache	30	3155	23	17.7
Non-White	Viral respiratory symptoms	30	3155	59.5	43.7
Non-White	Peripheral edema	30	3155	23.4	8.5
Non-White	Light headedness	30	3155	22.5	13.2
Non-White	Nausea	30	3155	25.5	14.5
Non-White	Palpitations	30	3155	9.9	5.6
Non-White	Seizure	30	3155	9	2.1
Non-White	Shortness of breath	30	3155	37.2	17.3
Non-White	Insomnia	30	3155	19.8	12.4
Non-White	Syncope	30	3155	10.6	3.9
Non-White	Bleeding	60	3155	20	9.5
Non-White	Chest pain	60	3155	29.5	16.8
Non-White	Diaphoresis	60	3155	4.3	2.7
Non-White	Headache	60	3155	21.6	17
Non-White	Viral respiratory symptoms	60	3155	57.2	41.6
Non-White	Peripheral edema	60	3155	23	7.9

Non-White	Light headedness	60	3155	22.1	12.5
Non-White	Nausea	60	3155	23.4	13.8
Non-White	Palpitations	60	3155	9.2	5.1
Non-White	Seizure	60	3155	8.1	1.9
Non-White	Shortness of breath	60	3155	35.4	16.2
Non-White	Insomnia	60	3155	18.7	11.7
Non-White	Syncope	60	3155	10.1	3.6
Non-White	Bleeding	90	3155	18.7	9.1
Non-White	Chest pain	90	3155	28.4	15.8
Non-White	Diaphoresis	90	3155	3.6	2.5
Non-White	Headache	90	3155	20.5	16
Non-White	Viral respiratory symptoms	90	3155	55.4	39.5
Non-White	Peripheral edema	90	3155	22.3	7.5
Non-White	Light headedness	90	3155	21.2	11.8
Non-White	Nausea	90	3155	22.3	13.1
Non-White	Palpitations	90	3155	9	4.8
Non-White	Seizure	90	3155	7.9	1.7
Non-White	Shortness of breath	90	3155	34.2	15.2
Non-White	Insomnia	90	3155	17.6	10.9
Non-White	Syncope	90	3155	9.9	3.4
Non-White	Bleeding	120	3155	18.5	8.7
Non-White	Chest pain	120	3155	26.1	14.9
Non-White	Diaphoresis	120	3155	3.2	2.3
Non-White	Headache	120	3155	19.4	15.2
Non-White	Viral respiratory symptoms	120	3155	54.1	37.7
Non-White	Peripheral edema	120	3155	20.9	7
Non-White	Light headedness	120	3155	20	11.4
Non-White	Nausea	120	3155	20.9	12.3
Non-White	Palpitations	120	3155	8.8	4.6
Non-White	Seizure	120	3155	7.7	1.5
Non-White	Shortness of breath	120	3155	33.6	14.2
Non-White	Insomnia	120	3155	17.1	10.5
Non-White	Syncope	120	3155	9.7	3.2
Non-White	Bleeding	150	3155	17.8	8.2
Non-White	Chest pain	150	3155	25	14.2
Non-White	Diaphoresis	150	3155	3.2	2.1

Non-White	Headache	150	3155	18.9	14.3
Non-White	Viral respiratory symptoms	150	3155	53.4	36.1
Non-White	Peripheral edema	150	3155	20.5	6.7
Non-White	Light headedness	150	3155	19.4	10.8
Non-White	Nausea	150	3155	20.7	11.6
Non-White	Palpitations	150	3155	8.3	4.3
Non-White	Seizure	150	3155	7.4	1.5
Non-White	Shortness of breath	150	3155	32.7	13.5
Non-White	Insomnia	150	3155	16.2	10.1
Non-White	Syncope	150	3155	9.2	3.1
Non-White	Bleeding	180	3155	16.9	7.6
Non-White	Chest pain	180	3155	24.5	13.6
Non-White	Diaphoresis	180	3155	3.2	2
Non-White	Headache	180	3155	18.5	14
Non-White	Viral respiratory symptoms	180	3155	51.8	34.8
Non-White	Peripheral edema	180	3155	19.8	6.2
Non-White	Light headedness	180	3155	18.9	10.4
Non-White	Nausea	180	3155	19.1	11.1
Non-White	Palpitations	180	3155	8.1	4
Non-White	Seizure	180	3155	7	1.4
Non-White	Shortness of breath	180	3155	31.5	12.9
Non-White	Insomnia	180	3155	15.8	9.7
Non-White	Syncope	180	3155	9	3
Non-White	Bleeding	210	3155	15.3	7.1
Non-White	Chest pain	210	3155	23.6	12.9
Non-White	Diaphoresis	210	3155	2.9	1.8
Non-White	Headache	210	3155	17.8	13.5
Non-White	Viral respiratory symptoms	210	3155	50.7	33.5
Non-White	Peripheral edema	210	3155	18.9	5.9
Non-White	Light headedness	210	3155	17.8	9.8
Non-White	Nausea	210	3155	17.6	10.3
Non-White	Palpitations	210	3155	7.9	3.8
Non-White	Seizure	210	3155	7	1.3
Non-White	Shortness of breath	210	3155	31.1	12.3
Non-White	Insomnia	210	3155	15.1	9.4
Non-White	Syncope	210	3155	8.6	2.7

Non-White	Bleeding	240	3155	14.9	6.8
Non-White	Chest pain	240	3155	23	12.4
Non-White	Diaphoresis	240	3155	2.7	1.7
Non-White	Headache	240	3155	16.7	13.2
Non-White	Viral respiratory symptoms	240	3155	49.1	32.4
Non-White	Peripheral edema	240	3155	18.2	5.5
Non-White	Light headedness	240	3155	16.9	9.5
Non-White	Nausea	240	3155	16.4	9.8
Non-White	Palpitations	240	3155	7.2	3.5
Non-White	Seizure	240	3155	6.3	1.3
Non-White	Shortness of breath	240	3155	28.8	11.6
Non-White	Insomnia	240	3155	14.2	9
Non-White	Syncope	240	3155	8.1	2.7
Non-White	Bleeding	270	3155	14.4	6.4
Non-White	Chest pain	270	3155	20.7	11.8
Non-White	Diaphoresis	270	3155	2.5	1.7
Non-White	Headache	270	3155	16	12.6
Non-White	Viral respiratory symptoms	270	3155	47.3	31.4
Non-White	Peripheral edema	270	3155	17.1	5.2
Non-White	Light headedness	270	3155	16	9.1
Non-White	Nausea	270	3155	15.1	9.1
Non-White	Palpitations	270	3155	6.8	3.4
Non-White	Seizure	270	3155	6.1	1.3
Non-White	Shortness of breath	270	3155	27	11.1
Non-White	Insomnia	270	3155	13.5	8.4
Non-White	Syncope	270	3155	7.7	2.7
Non-White	Bleeding	300	3155	13.3	6.1
Non-White	Chest pain	300	3155	18.9	10.9
Non-White	Diaphoresis	300	3155	2.3	1.7
Non-White	Headache	300	3155	14.9	12
Non-White	Viral respiratory symptoms	300	3155	45.7	30.2
Non-White	Peripheral edema	300	3155	16	5
Non-White	Light headedness	300	3155	15.3	8.7
Non-White	Nausea	300	3155	14.2	8.6
Non-White	Palpitations	300	3155	6.5	3.3
Non-White	Seizure	300	3155	5.9	1.3

Non-White	Shortness of breath	300	3155	25.9	10.4
Non-White	Insomnia	300	3155	13.3	8
Non-White	Syncope	300	3155	7.4	2.5
Non-White	Bleeding	330	3155	12.6	5.9
Non-White	Chest pain	330	3155	18.7	10.1
Non-White	Diaphoresis	330	3155	2.3	1.5
Non-White	Headache	330	3155	14	11.1
Non-White	Viral respiratory symptoms	330	3155	43.5	28.7
Non-White	Peripheral edema	330	3155	15.3	4.6
Non-White	Light headedness	330	3155	14.6	8.3
Non-White	Nausea	330	3155	14	8
Non-White	Palpitations	330	3155	6.1	3.2
Non-White	Seizure	330	3155	5.6	1.2
Non-White	Shortness of breath	330	3155	25.5	9.8
Non-White	Insomnia	330	3155	12.6	7.6
Non-White	Syncope	330	3155	7	2.4
Non-White	Bleeding	360	3155	11.9	5.4
Non-White	Chest pain	360	3155	17.3	9.7
Non-White	Diaphoresis	360	3155	2	1.5
Non-White	Headache	360	3155	13.5	10.3
Non-White	Viral respiratory symptoms	360	3155	42.8	27.6
Non-White	Peripheral edema	360	3155	14.6	4.4
Non-White	Light headedness	360	3155	13.3	7.9
Non-White	Nausea	360	3155	13.7	7.5
Non-White	Palpitations	360	3155	5.9	2.9
Non-White	Seizure	360	3155	5.4	1.1
Non-White	Shortness of breath	360	3155	23.9	9.3
Non-White	Insomnia	360	3155	11.3	7.1
Non-White	Syncope	360	3155	6.8	2.1
Established CVD	Bleeding	30	3450	26.8	18.9
Established CVD	Chest pain	30	3450	39.7	37
Established CVD	Diaphoresis	30	3450	5.5	3.4
Established CVD	Headache	30	3450	23.3	20.1
Established CVD	Viral respiratory symptoms	30	3450	67.1	59.2
Established CVD	Peripheral edema	30	3450	35.3	23.4
Established CVD	Light headedness	30	3450	30.8	23.2

Established CVD	Nausea	30	3450	29.9	20.5
Established CVD	Palpitations	30	3450	15	12.4
Established CVD	Seizure	30	3450	8.4	3.4
Established CVD	Shortness of breath	30	3450	50.1	38.5
Established CVD	Insomnia	30	3450	23.9	17.9
Established CVD	Syncope	30	3450	13.1	9.5
Established CVD	Bleeding	60	3450	25.4	17.8
Established CVD	Chest pain	60	3450	38.3	34.7
Established CVD	Diaphoresis	60	3450	5.2	3.2
Established CVD	Headache	60	3450	22.3	19.2
Established CVD	Viral respiratory symptoms	60	3450	63.7	56.2
Established CVD	Peripheral edema	60	3450	33.5	22
Established CVD	Light headedness	60	3450	29.8	22.2
Established CVD	Nausea	60	3450	27.4	19.5
Established CVD	Palpitations	60	3450	13.7	11.7
Established CVD	Seizure	60	3450	7.7	3.2
Established CVD	Shortness of breath	60	3450	47.6	36
Established CVD	Insomnia	60	3450	22.1	17
Established CVD	Syncope	60	3450	12.2	9
Established CVD	Bleeding	90	3450	24	16.9
Established CVD	Chest pain	90	3450	37	32.7
Established CVD	Diaphoresis	90	3450	4.9	3.1
Established CVD	Headache	90	3450	21.1	18.4
Established CVD	Viral respiratory symptoms	90	3450	60.9	53.9
Established CVD	Peripheral edema	90	3450	31.6	20.9
Established CVD	Light headedness	90	3450	28.8	20.9
Established CVD	Nausea	90	3450	26.5	18.6
Established CVD	Palpitations	90	3450	13.3	11.1
Established CVD	Seizure	90	3450	7.4	2.9
Established CVD	Shortness of breath	90	3450	46	34.2
Established CVD	Insomnia	90	3450	21.1	16.1
Established CVD	Syncope	90	3450	11.5	8.7
Established CVD	Bleeding	120	3450	23.5	16.1
Established CVD	Chest pain	120	3450	35.3	31.2
Established CVD	Diaphoresis	120	3450	4.6	3
Established CVD	Headache	120	3450	20.2	17.5

Established CVD	Viral respiratory symptoms	120	3450	59.4	51.6
Established CVD	Peripheral edema	120	3450	31.1	19.8
Established CVD	Light headedness	120	3450	26.1	20.1
Established CVD	Nausea	120	3450	25.7	17.4
Established CVD	Palpitations	120	3450	13.1	10.7
Established CVD	Seizure	120	3450	6.8	2.7
Established CVD	Shortness of breath	120	3450	44.2	32.6
Established CVD	Insomnia	120	3450	20.8	15.4
Established CVD	Syncope	120	3450	11.1	8.2
Established CVD	Bleeding	150	3450	22.9	15.3
Established CVD	Chest pain	150	3450	33.8	29.8
Established CVD	Diaphoresis	150	3450	4.6	2.7
Established CVD	Headache	150	3450	19.3	16.6
Established CVD	Viral respiratory symptoms	150	3450	57.8	49.6
Established CVD	Peripheral edema	150	3450	29.9	18.6
Established CVD	Light headedness	150	3450	25.5	19
Established CVD	Nausea	150	3450	25.1	16.5
Established CVD	Palpitations	150	3450	12.2	10.2
Established CVD	Seizure	150	3450	6.8	2.6
Established CVD	Shortness of breath	150	3450	42.2	31.2
Established CVD	Insomnia	150	3450	20.2	14.6
Established CVD	Syncope	150	3450	10.9	7.9
Established CVD	Bleeding	180	3450	21.8	14.4
Established CVD	Chest pain	180	3450	32.2	28.4
Established CVD	Diaphoresis	180	3450	4.4	2.6
Established CVD	Headache	180	3450	19.3	16
Established CVD	Viral respiratory symptoms	180	3450	57.1	47.5
Established CVD	Peripheral edema	180	3450	28.9	17.5
Established CVD	Light headedness	180	3450	25.1	18.1
Established CVD	Nausea	180	3450	24.8	15.4
Established CVD	Palpitations	180	3450	11.7	9.5
Established CVD	Seizure	180	3450	6.5	2.4
Established CVD	Shortness of breath	180	3450	40.7	30
Established CVD	Insomnia	180	3450	19.6	13.8
Established CVD	Syncope	180	3450	10.3	7.6
Established CVD	Bleeding	210	3450	20.6	13.5

Established CVD	Chest pain	210	3450	30.5	27.1
Established CVD	Diaphoresis	210	3450	4.4	2.5
Established CVD	Headache	210	3450	18.4	15.4
Established CVD	Viral respiratory symptoms	210	3450	55.5	45.9
Established CVD	Peripheral edema	210	3450	27.7	16.5
Established CVD	Light headedness	210	3450	23.6	17.7
Established CVD	Nausea	210	3450	23.2	14.7
Established CVD	Palpitations	210	3450	11.2	9.1
Established CVD	Seizure	210	3450	6.2	2.3
Established CVD	Shortness of breath	210	3450	39.5	28.2
Established CVD	Insomnia	210	3450	18.7	13.2
Established CVD	Syncope	210	3450	9.9	7.2
Established CVD	Bleeding	240	3450	20.1	12.5
Established CVD	Chest pain	240	3450	29.2	25.8
Established CVD	Diaphoresis	240	3450	4.1	2.3
Established CVD	Headache	240	3450	17.8	14.8
Established CVD	Viral respiratory symptoms	240	3450	54.1	44.1
Established CVD	Peripheral edema	240	3450	26.5	15.5
Established CVD	Light headedness	240	3450	22.1	16.9
Established CVD	Nausea	240	3450	21.8	14
Established CVD	Palpitations	240	3450	10.5	8.8
Established CVD	Seizure	240	3450	5.6	2.2
Established CVD	Shortness of breath	240	3450	37.9	27.1
Established CVD	Insomnia	240	3450	18	12.4
Established CVD	Syncope	240	3450	9.6	6.9
Established CVD	Bleeding	270	3450	19	11.9
Established CVD	Chest pain	270	3450	27.6	24.7
Established CVD	Diaphoresis	270	3450	4	2.2
Established CVD	Headache	270	3450	17.1	14
Established CVD	Viral respiratory symptoms	270	3450	52.7	42.6
Established CVD	Peripheral edema	270	3450	25.4	14.2
Established CVD	Light headedness	270	3450	21.5	16.3
Established CVD	Nausea	270	3450	20.5	13.2
Established CVD	Palpitations	270	3450	10.2	8.5
Established CVD	Seizure	270	3450	5.5	2.1
Established CVD	Shortness of breath	270	3450	36.9	26

Established CVD	Insomnia	270	3450	17.6	11.9
Established CVD	Syncope	270	3450	9.1	6.6
Established CVD	Bleeding	300	3450	17.7	11.5
Established CVD	Chest pain	300	3450	26.1	23.9
Established CVD	Diaphoresis	300	3450	3.7	2
Established CVD	Headache	300	3450	16.4	13.4
Established CVD	Viral respiratory symptoms	300	3450	51	41.1
Established CVD	Peripheral edema	300	3450	23.7	13.6
Established CVD	Light headedness	300	3450	19.6	15.5
Established CVD	Nausea	300	3450	19.5	12.4
Established CVD	Palpitations	300	3450	9.7	8.2
Established CVD	Seizure	300	3450	5.2	2.1
Established CVD	Shortness of breath	300	3450	35	24.9
Established CVD	Insomnia	300	3450	16.8	11.5
Established CVD	Syncope	300	3450	8.7	6.3
Established CVD	Bleeding	330	3450	16.4	10.7
Established CVD	Chest pain	330	3450	25.7	22.5
Established CVD	Diaphoresis	330	3450	3.5	1.8
Established CVD	Headache	330	3450	15.9	12.7
Established CVD	Viral respiratory symptoms	330	3450	50	39.7
Established CVD	Peripheral edema	330	3450	22.7	12.7
Established CVD	Light headedness	330	3450	19	15
Established CVD	Nausea	330	3450	19	11.6
Established CVD	Palpitations	330	3450	8.8	7.6
Established CVD	Seizure	330	3450	5	2.1
Established CVD	Shortness of breath	330	3450	34.1	23.7
Established CVD	Insomnia	330	3450	15.9	10.8
Established CVD	Syncope	330	3450	8.3	5.9
Established CVD	Bleeding	360	3450	15.2	10.2
Established CVD	Chest pain	360	3450	23.3	21.8
Established CVD	Diaphoresis	360	3450	3.4	1.7
Established CVD	Headache	360	3450	14.9	12.1
Established CVD	Viral respiratory symptoms	360	3450	47.2	38.5
Established CVD	Peripheral edema	360	3450	21.2	12
Established CVD	Light headedness	360	3450	17.7	14.1
Established CVD	Nausea	360	3450	17.7	11

Established CVD	Palpitations	360	3450	8.4	7
Established CVD	Seizure	360	3450	4.4	1.9
Established CVD	Shortness of breath	360	3450	32.3	22.5
Established CVD	Insomnia	360	3450	14.9	10.2
Established CVD	Syncope	360	3450	8	5.8
Non-established CVD	Bleeding	30	15227	14.1	8.2
Non-established CVD	Chest pain	30	15227	18.1	11.9
Non-established CVD	Diaphoresis	30	15227	4.1	1.9
Non-established CVD	Headache	30	15227	19.7	15.3
Non-established CVD	Viral respiratory symptoms	30	15227	52	37.7
Non-established CVD	Peripheral edema	30	15227	12.9	6
Non-established CVD	Light headedness	30	15227	13.4	9.7
Non-established CVD	Nausea	30	15227	20.6	12.6
Non-established CVD	Palpitations	30	15227	5.3	4.2
Non-established CVD	Seizure	30	15227	6.4	2.1
Non-established CVD	Shortness of breath	30	15227	24.6	12.4
Non-established CVD	Insomnia	30	15227	18.3	11.3
Non-established CVD	Syncope	30	15227	5.5	2.9
Non-established CVD	Bleeding	60	15227	13.3	7.4
Non-established CVD	Chest pain	60	15227	17	10.9
Non-established CVD	Diaphoresis	60	15227	3.9	1.8
Non-established CVD	Headache	60	15227	18	14.4
Non-established CVD	Viral respiratory symptoms	60	15227	49.4	35.1
Non-established CVD	Peripheral edema	60	15227	12.2	5.5
Non-established CVD	Light headedness	60	15227	12.5	9
Non-established CVD	Nausea	60	15227	19.1	11.7

Non-established CVD	Palpitations	60	15227	4.9	3.8
Non-established CVD	Seizure	60	15227	6	1.9
Non-established CVD	Shortness of breath	60	15227	22.7	11.6
Non-established CVD	Insomnia	60	15227	16.8	10.5
Non-established CVD	Syncope	60	15227	5	2.7
Non-established CVD	Bleeding	90	15227	12.5	6.9
Non-established CVD	Chest pain	90	15227	16	10.2
Non-established CVD	Diaphoresis	90	15227	3.7	1.7
Non-established CVD	Headache	90	15227	17	13.5
Non-established CVD	Viral respiratory symptoms	90	15227	46.3	33
Non-established CVD	Peripheral edema	90	15227	12	5.1
Non-established CVD	Light headedness	90	15227	12	8.3
Non-established CVD	Nausea	90	15227	17.9	10.8
Non-established CVD	Palpitations	90	15227	4.7	3.6
Non-established CVD	Seizure	90	15227	5.9	1.8
Non-established CVD	Shortness of breath	90	15227	21.4	10.7
Non-established CVD	Insomnia	90	15227	15.9	9.9
Non-established CVD	Syncope	90	15227	4.3	2.4
Non-established CVD	Bleeding	120	15227	11.4	6.6
Non-established CVD	Chest pain	120	15227	15	9.6
Non-established CVD	Diaphoresis	120	15227	3.7	1.6
Non-established CVD	Headache	120	15227	16.2	12.8
Non-established CVD	Viral respiratory symptoms	120	15227	44.4	31.2
Non-established CVD	Peripheral edema	120	15227	10.9	4.7
Non-established CVD	Light headedness	120	15227	11.6	7.9

Non-established CVD	Nausea	120	15227	16.7	10.1
Non-established CVD	Palpitations	120	15227	4.5	3.4
Non-established CVD	Seizure	120	15227	5.9	1.7
Non-established CVD	Shortness of breath	120	15227	19.8	10
Non-established CVD	Insomnia	120	15227	15.5	9.3
Non-established CVD	Syncope	120	15227	4.2	2.3
Non-established CVD	Bleeding	150	15227	10.9	6.2
Non-established CVD	Chest pain	150	15227	14.3	9.2
Non-established CVD	Diaphoresis	150	15227	3.5	1.5
Non-established CVD	Headache	150	15227	15.5	12.1
Non-established CVD	Viral respiratory symptoms	150	15227	43.2	29.7
Non-established CVD	Peripheral edema	150	15227	10.4	4.4
Non-established CVD	Light headedness	150	15227	10.9	7.4
Non-established CVD	Nausea	150	15227	16	9.5
Non-established CVD	Palpitations	150	15227	4.5	3.3
Non-established CVD	Seizure	150	15227	5.8	1.6
Non-established CVD	Shortness of breath	150	15227	18.8	9.4
Non-established CVD	Insomnia	150	15227	14.8	8.9
Non-established CVD	Syncope	150	15227	4.1	2.1
Non-established CVD	Bleeding	180	15227	10.2	5.8
Non-established CVD	Chest pain	180	15227	13.9	8.7
Non-established CVD	Diaphoresis	180	15227	3.5	1.4
Non-established CVD	Headache	180	15227	15	11.5
Non-established CVD	Viral respiratory symptoms	180	15227	41.7	28.2
Non-established CVD	Peripheral edema	180	15227	10	4

Non-established CVD	Light headedness	180	15227	10.6	7
Non-established CVD	Nausea	180	15227	14.6	8.9
Non-established CVD	Palpitations	180	15227	4.5	3.1
Non-established CVD	Seizure	180	15227	5.5	1.5
Non-established CVD	Shortness of breath	180	15227	18.1	8.8
Non-established CVD	Insomnia	180	15227	13.8	8.5
Non-established CVD	Syncope	180	15227	3.5	2
Non-established CVD	Bleeding	210	15227	9.1	5.5
Non-established CVD	Chest pain	210	15227	13.1	8.2
Non-established CVD	Diaphoresis	210	15227	3.2	1.4
Non-established CVD	Headache	210	15227	14.1	10.9
Non-established CVD	Viral respiratory symptoms	210	15227	39	26.8
Non-established CVD	Peripheral edema	210	15227	9.5	3.7
Non-established CVD	Light headedness	210	15227	9.6	6.5
Non-established CVD	Nausea	210	15227	13.1	8.2
Non-established CVD	Palpitations	210	15227	4.2	2.9
Non-established CVD	Seizure	210	15227	5.1	1.4
Non-established CVD	Shortness of breath	210	15227	17.2	8.3
Non-established CVD	Insomnia	210	15227	13.3	8.1
Non-established CVD	Syncope	210	15227	3.5	1.9
Non-established CVD	Bleeding	240	15227	8.7	5.1
Non-established CVD	Chest pain	240	15227	12.5	7.8
Non-established CVD	Diaphoresis	240	15227	3.2	1.3
Non-established CVD	Headache	240	15227	13.3	10.3
Non-established CVD	Viral respiratory symptoms	240	15227	37.2	25.6

Non-established CVD	Peripheral edema	240	15227	8.7	3.5
Non-established CVD	Light headedness	240	15227	9.1	6.1
Non-established CVD	Nausea	240	15227	12.2	7.7
Non-established CVD	Palpitations	240	15227	4.1	2.8
Non-established CVD	Seizure	240	15227	4.7	1.4
Non-established CVD	Shortness of breath	240	15227	15.9	7.9
Non-established CVD	Insomnia	240	15227	12.5	7.7
Non-established CVD	Syncope	240	15227	3.4	1.8
Non-established CVD	Bleeding	270	15227	8.5	4.7
Non-established CVD	Chest pain	270	15227	11.6	7.5
Non-established CVD	Diaphoresis	270	15227	2.8	1.2
Non-established CVD	Headache	270	15227	12.5	9.8
Non-established CVD	Viral respiratory symptoms	270	15227	35.6	24.4
Non-established CVD	Peripheral edema	270	15227	8.3	3.2
Non-established CVD	Light headedness	270	15227	8.5	5.8
Non-established CVD	Nausea	270	15227	11.6	7.1
Non-established CVD	Palpitations	270	15227	4.1	2.7
Non-established CVD	Seizure	270	15227	4.2	1.3
Non-established CVD	Shortness of breath	270	15227	14.8	7.4
Non-established CVD	Insomnia	270	15227	11.8	7.2
Non-established CVD	Syncope	270	15227	3.2	1.7
Non-established CVD	Bleeding	300	15227	7.9	4.5
Non-established CVD	Chest pain	300	15227	10.6	7
Non-established CVD	Diaphoresis	300	15227	2.6	1.2
Non-established CVD	Headache	300	15227	11.8	9.3

Non-established CVD	Viral respiratory symptoms	300	15227	33.5	23.3
Non-established CVD	Peripheral edema	300	15227	7.4	3
Non-established CVD	Light headedness	300	15227	8.4	5.5
Non-established CVD	Nausea	300	15227	10.5	6.8
Non-established CVD	Palpitations	300	15227	3.9	2.5
Non-established CVD	Seizure	300	15227	3.7	1.3
Non-established CVD	Shortness of breath	300	15227	14.2	7
Non-established CVD	Insomnia	300	15227	11.2	6.8
Non-established CVD	Syncope	300	15227	2.9	1.7
Non-established CVD	Bleeding	330	15227	7.6	4.2
Non-established CVD	Chest pain	330	15227	10.2	6.5
Non-established CVD	Diaphoresis	330	15227	2.2	1.1
Non-established CVD	Headache	330	15227	11.3	8.8
Non-established CVD	Viral respiratory symptoms	330	15227	31.5	22.1
Non-established CVD	Peripheral edema	330	15227	7.1	2.8
Non-established CVD	Light headedness	330	15227	8.1	5.3
Non-established CVD	Nausea	330	15227	10.1	6.4
Non-established CVD	Palpitations	330	15227	3.7	2.3
Non-established CVD	Seizure	330	15227	3.4	1.2
Non-established CVD	Shortness of breath	330	15227	13.3	6.5
Non-established CVD	Insomnia	330	15227	10.5	6.4
Non-established CVD	Syncope	330	15227	2.6	1.6
Non-established CVD	Bleeding	360	15227	7.6	3.8
Non-established CVD	Chest pain	360	15227	10	6.1
Non-established CVD	Diaphoresis	360	15227	2	1

Non-established CVD	Headache	360	15227	10.6	8.2
Non-established CVD	Viral respiratory symptoms	360	15227	30.9	21
Non-established CVD	Peripheral edema	360	15227	6.3	2.6
Non-established CVD	Light headedness	360	15227	7.5	5
Non-established CVD	Nausea	360	15227	9.6	5.8
Non-established CVD	Palpitations	360	15227	3.7	2.2
Non-established CVD	Seizure	360	15227	3.4	1.1
Non-established CVD	Shortness of breath	360	15227	12.7	6
Non-established CVD	Insomnia	360	15227	9.6	6
Non-established CVD	Syncope	360	15227	2.4	1.4

Table 3. Primary care clinics selected.

Primary care clinic name
HMC ADULT MEDICINE CLINIC 20
HMC AFTERCARE CLINIC
HMC FAMILY MEDICAL WELLNESS CLINIC
HMC FAMILY MEDICINE CLINIC 20
HMC INTERNATIONAL MEDICINE CLINIC 20
HMC NWFC PEDIATRIC CLINIC
HMC PEDIATRIC CLINICS 20
U WHCC ARNP
U WHCC GEN INT MED
UW MED PRIMARY CARE
UWMC CAMPUS HEALTH SERV
UWMC GENERAL IM
UWMPC KIRKLAND FAMILY MEDICINE
UWMPC VIRTUAL PRIMARY CARE
HMC PSQ - 3RD AVE CENTER 20
HMC PSQ FIELD BASED PRIMARY CARE CLINIC 20
HMC PSQ-PIONEER SQUARE CLINIC 20
HMC SENIOR CARE CLINIC 20
HMCPB FAMILY MEDICAL CLINIC

LINK BANNER UNIVERSITY MEDICINE-PRIMARY CARE CLINIC-MARANA
LINK BRIDGE PRIMARY CARE
LINK CAPITAL EASTSIDE FAMILY PRACTICE
LINK IORA PRIMARY CARE-CENTRAL DISTRICT
LINK ISLAND PRIMARY CARE ORCAS
LINK SOUTHCENTRAL FOUNDATION PRIMARY CARE CENTERS
LINK UWM COSMAS PRIMARY CARE
LINK UWM EQUINOX PRIMARY CARE
LINK UWM EVERGREEN PRIMARY CARE SULTAN
LINK UWM GREENLAKE PRIMARY CARE - INTERNAL MEDICINE/PEDIATRICS
LINK UWM IORA PRIMARY CARE
LINK UWM JEFFERSON HEALTHCARE PRIMARY CARE
LINK UWM PEACEHEALTH PRIMARY CARE - KETCHIKAN
LINK UWM PETER SHALIT MD - PRIMARY CARE
LINK UWM SOUND PRIMARY CARE
LINK UWM SWEDISH PRIMARY CARE
LINK UWM SWEDISH SANDPOINT PRIMARY CARE
UWPC BALLARD FAMILY MEDICINE
UWPC BELLTOWN FAMILY MEDICINE
UWPC BELLTOWN INTERNAL MEDICINE
UWPC FACTORIA FAMILY MEDICINE
UWPC FACTORIA INTERNAL MEDICINE
UWPC FEDERAL WAY FAMILY MEDICINE
UWPC FEDERAL WAY INTERNAL MEDICINE
UWPC FREMONT PRIMARY CARE
UWPC ISSAQUAH FAMILY MEDICINE
UWPC KDM FAMILY MEDICINE
UWPC KDM INTERNAL MEDICINE
UWPC KDM PEDIATRIC MEDICINE
UWPC LOPEZ ISLAND FAMILY MEDICINE
UWPC MOUNTLAKE TERRACE FAMILY MEDICINE
UWPC MOUNTLAKE TERRACE INTERNAL MEDICINE
UWPC NORTHGATE FAMILY MEDICINE
UWPC NURSING SERVICES
UWPC OPMC PRIMARY CARE
UWPC RAVENNA FAMILY MEDICINE
UWPC SHORELINE FAMILY MEDICINE
UWPC SHORELINE INTERNAL MEDICINE

UWPC SHORELINE PEDS
UWPC SLU FAMILY MEDICINE
UWPC WOODINVILLE FAMILY MEDICINE
UWPC WOODINVILLE INTERNAL MEDICINE
ZUWMC FAMILY MED CENTER

Table 4. Visit types used.

Visit type
RETURN OFFSITE
RETURN
TELEMEDICINE RETURN
NEW ACUTE
PHONE SHORT
PHONE VISIT
TELEMEDICINE NEW
COMPLEX
ECARE TELEMEDICINE REMOTE
RETURN WELLNESS
RETURN MEDICARE AWV
TELEMEDICINE RETURN LONG
RETURN LONG
NEW
PHONE LONG
NEW WELLNESS
PRIM/PREP RETURN
TELEMEDICINE MEDICARE AWV RET
NEW OB
TELEMEDICINE GROUP RETURN
RET OB
PRIM/PREP NEW
PROCEDURE
ED FOLLOW-UP
NEW MEDICARE AWV
RETURN OCCUPATIONAL HLTH LONG
PRE-OP
NEW OFFSITE
E-VISIT

RETURN OCCUPATIONAL HEALTH
CONSULT
TELEMEDICINE GROUP NEW
POST-OP
WALK IN
RETURN SUBSPECIALTY
TELEMEDICINE CLINICAL SUPPORT
POST PARTUM
NEW SUBSPECIALTY
NST
NEW PROCEDURE
ACUPUNCTURE
ACUPUNCTURE NEW
MYCHART ONLY - PC RETURN
PHYSICAL
PEP NEW
LACTATION CONSULT
PATIENT OUTREACH
POST PARTUM SHORT
ACUTE VISIT
UOBTV
NEW OCCUPATIONAL HEALTH
UOBL
PEP RETURN OCCUPATIONAL
PEP RETURN
PEP NEW OCCUPATIONAL
RETURN RSV PEDS
UTVAG
TELEMEDICINE MPX TREATMENT
WOUND CARE

Table 5. Long COVID specialist derived symptom mapping for CUIs and ICD codes.

Symptom	CUI	ICD	Source
Abdominal pain	abdominal pain	789.0,789.00,R10.9	Specialist
Bloating	abdominal bloating	R14.0	Specialist
Breathing faster than normal	tachypnea	786.06,R06.82	Specialist
Confusion, difficulty thinking	confusion	R41.0	Specialist

Diarrhea	diarrhea	787.91,R19.7	Specialist
Difficulty hearing	hearing impairment	389.9,389,H91.9	Specialist
Difficulty sleeping (too much, too little, early awakening)	non-organic sleep disorder, sleeplessness, snore, sleep disorders	307.40,780.52,G47.00,G47.0,R06.83,G47.9,G47	Specialist
Difficulty swallowing	deglutition disorders	787.20,787.2,R13.10,R13.1	Specialist
Fainting or blackouts	syncope	780.2,R55	Specialist
General/muscle weakness	generalized muscle weakness	728.87,M62.81	Specialist
Heart palpitations, pulse skips, heart block	palpitations, premature ventricular contractions	785.1,R00.2,I49.3	Specialist
Indigestion or esophageal/"acid" reflux	gastroesophageal reflux disease	530.81,K21,K21.9	Specialist
Joint pain or swelling	arthralgia,polyarthritis, hydrarthrosis	719.4,719.40,714.9,719.00,719.0,M25.4,M25.5,M06.4,M13.0,M05,M06,M07,M08,M09,M10,M11,M12,M13,M14	Specialist
Loss of smell	anosmia	R43.0	Specialist
Loss of taste	dysgeusia	R43.2	Specialist
Memory problems or forgetfulness	memory loss, mild cognitive disorder, "mild cognitive impairment, so stated", amnesia	780.93,331.83,G31.84,R41.3	Specialist
Mood swings, irritability, depression	irritable mood, mental depression, mood swings,recurrent major depressive episodes, irritability and anger	799.22,799.24,296.3,296.30,F33.9,R45.86,F33,R45.4	Specialist
Muscle pain or cramps	myalgia, cramp and spasm	M62.831,M79.1,R25.2	Specialist
Nausea	nausea	R11.0	Specialist
Problems seeing (double or blurry vision)	diplopia	368.2,H53.2	Specialist
Rapid heart rate (Tachycardia)	tachycardia	785.0,R00.0	Specialist
Sexual dysfunction or loss of libido	erectile dysfunction, sexual dysfunction	F52.21,N52.9,N52,R37,F52.9	Specialist
Skin lesions (rash or lumpy lesions)	skin rash	782.1,R21	Specialist
Sore throat	sore throat	784.1,R07.0	Specialist
Swelling	edema	782.3,R60.9	Specialist
Tingling, numbness, burning, stabbing or "pins and needles"	tingling of skin, neuropathy	R20.2,G62.9	Specialist
Tinnitus (ringing in the ears)	tinnitus	388.30,388.3,H93.19,H93.1	Specialist

Unexplained hair loss	nonscarring hair loss, unspecified	L65.9	Specialist
Unexplained menstrual irregularity	menorrhagia, irregular menstruation (diagnosis)	626.4,N92.0,N92.6	Specialist
Urinary incontinence or difficulty urinating	difficulty passing urine, urinary hesitation	788.64,R30.0,R39.11	Specialist
Vomiting	vomiting	R11.10,R11.1	Specialist
brain fog	unspecified symptoms and signs involving cognitive functions and awareness, other symptoms and signs involving cognitive functions and awareness	R41.9,R41.89,R41.8	Specialist
difficulty concentration	other specified cognitive deficit, attention and concentration deficit	R41.84,R41.840	Specialist
malaise	malaise	R53.81	Specialist
word-finding difficulty	aphasia	784.3,R47.01	Specialist
Cough		786.2,491.0,R05	Literature
Dizziness		780.4,R42	Literature
Fatigue or tiredness		780.79,R53.81,R53.8,R53.83,R53.1	Literature
Headache		784.0,R51	Literature
Lack of appetite		783,R63.0	Literature
Shortness of breath		786.05,786.0,786.9,R06.02,R06.00,R06.09	Literature
Weight loss		783.21,R63.4	Literature
chest pain		786.5,786.50,786.51,786.52,786.59,R07.9,R07.81	Literature
fever		780.6,780.60,R50.9	Literature

Table 6. NLP symptom to synonym mapping.

NLP symptom	Synonym
altered mental status	ams, confused, confusion
anxiety	agitated, agitation, anxious
arthralgia	arthralgias
bleeding	bleed, blood, bloody
bruising	bruise, bruises, ecchymosis
chest pain	cp
chills	chill
cough	c, c., cough cough, coughing, coughs, distress coughing, distressed coughing
cramping	cramps
decreased appetite	loss of appetite, poor appetite, poor p.o. intake, poor po intake, reduced appetite
deformities	deformity

dehydration	dehydrated
diarrhea	d, d., diarrhea stools, loose stools
disharge	drainage
distended	distention
dysphagia	difficulty swallowing, dysphagia symptoms
erythema	erythematous, redness
exudates	exudate
fall	falls
fatigue	drowsiness, drowsy, fatigued, somnolence, somnolent, tired, tiredness
fever	f, f., febrile, fevers
flu-like symptoms	flu - like symptoms, influenza - like symptoms
gi symptoms	abdominal symptoms
headache	ha, headaches
heartburn	gerd symptoms, heartburn symptoms
hematochezia	brbpr
ill	ill - appearing, ill appearing, ill symptoms, illness, sick
incontinent	incontinence
irritation	irritable
itching	itchy
lethargy	lethargic
lightheadedness	dizziness, dizzy, headedness, lightheaded
myalgia	ache, aches, aching, bodyaches, myalgias
nausea	n, n., nauseated, nauseous
pain	discomfort, painful, pains
pruritus	pruritis
rash	rashes
respiratory symptoms	uri symptoms
runny nose	rhinorrhea
seizures	seizure, seizures
shortness of breath	shortness of breath, difficult breathing, difficulty breathing, difficulty of breathing, distress breathing, distressed breathing, doe, dsypnea, dypsnea, dyspnea, dyspnea exertion, dyspnea on exertion, increase work of breathing, increased work of breathing, out of breath, respiratory distress, short of breath, shortneses of breath, shortness breath, shortness of breaths, sob, sob on exertion, trouble breathing, work of breathing
sore throat	pharyngitis
soreness	sore
sputum	sputum production

sweats	diaphoresis, nightsweats, sweating
swelling	edema, oedema, swollen
syncope	fainting
tenderness	tender
tremors	tremor
ulcers	ulcer, ulceration, ulcerations
urinary	symptoms urinary
urination	urinating
vomiting	emesis, v, v., vomitting
weakness	weak
wheezing	wheeze, wheezes
wounds	wound

Table 7. Potential Long COVID symptom to NLP symptom mapping.

Potential Long COVID symptom	NLP symptom
Fatigue or tiredness	fatigue
Fatigue or tiredness	weakness
brain fog	altered mental status
brain fog	confusion
Shortness of breath	shortness of breath
Shortness of breath	respiratory symptoms
Difficulty sleeping (too much, too little, early awakening)	anxiety
Difficulty sleeping (too much, too little, early awakening)	lethargy
malaise	ill
Dizziness	lightheadedness
Rapid heart rate (Tachycardia)	heartburn
Memory problems or forgetfulness	altered mental status
Indigestion or esophageal/"acid" reflux	heartburn
difficulty concentration	altered mental status
chest pain	chest pain
Diarrhea	diarrhea
Nausea	nausea
Skin lesions (rash or lumpy lesions)	rash
Skin lesions (rash or lumpy lesions)	erythema
Skin lesions (rash or lumpy lesions)	pruritus
Abdominal pain	gi symptoms
Bloating	gi symptoms

General/muscle weakness	weakness
Mood swings, irritability, depression	anxiety
Mood swings, irritability, depression	lethargy
Muscle pain or cramps	myalgia
Muscle pain or cramps	cramping
Difficulty swallowing	dysphagia
Fainting or blackouts	syncope
Swelling	swelling
fever	fever
Confusion, difficulty thinking	altered mental status
Joint pain or swelling	arthralgia
Joint pain or swelling	swelling
Lack of appetite	decreased appetite
Sore throat	sore throat
Vomiting	vomiting
Breathing faster than normal	shortness of breath
Cough	cough
Headache	headache
Urinary incontinence or difficulty urinating	incontinent
Urinary incontinence or difficulty urinating	urinary

Table 8. Hyperparameters for model selection.

Model	Hyperparameter	Values
KNN	number of neighbors	3, 5, 7, 9, 11
	weights	uniform, distance
	metric	Euclidean, Manhattan
logistic regression	penalty	L1, L2
	C	0.001, 0.01, 0.1, 1.0, 10.0
	class weight	none, balanced
neural network	hidden layer sizes	(50), (100), (50, 50), (100, 50)
	activation function	relu, tanh
	alpha	0.0001, 0.001, 0.01
	learning rate	constant, inverse scaling, adaptive
	max iteration	200, 400, 600
random forest	number of estimators	50, 100, 200
	max depth	none, 10, 20, 30

	min sample split	2, 5, 10
	min sample leaf	1, 2, 4
	max features	auto, sqrt, log2
	class weight	none, balanced
XGBoost	learning rate	0.01, 0.1, 0.2
	number of estimators	50, 100, 200
	max depth	3, 5, 7
	subsample	0.8, 1.0
	subsample ratio of columns by tree	0.8, 1.0

VITA

Weipeng's research explores the intersection of natural language processing (NLP) and healthcare. He studies clinical note section classification and symptom extraction. He also used clinical narratives to gain insights about out-of-hospital cardiac arrest and Long COVID. He previously earned an undergraduate degree in Computer Science from the University of Wisconsin.