ABSTRACT

ESSAYS ON HEALTH CARE QUALITY: TIMELINESS, EQUITY, AND EFFICIENCY

by
Abubakar-Sadiq Bouda Abdulai

According to the National Academy of Medicine (NAM) (formerly called the Institute of Medicine), a quality health care system embodies six attributes: *timeliness, equity, safety, efficiency, effectiveness,* and *patient-centeredness.* Timeliness is to avoid unnecessary delays in care delivery for patients and caregivers; equity is to ensure that the quality of care that patients receive does not vary based on their personal characteristics; safety is to ensure that the care that is intended to help patients does not harm them; efficiency is to avoid waste and optimize resource allocation to improve care delivery; effective care is one that relies on sound scientific knowledge and delivers the most benefit to the patient; and patient-centeredness concerns the contributions of patients, their family members, and care givers to the patient's health. Thus, to make progress in quality care improvement, every aspect of the health care system as related to these six items must be improved.

Recent efforts targeted at improving quality of care in urgent care settings have included benchmarking and performance measurement, financial incentivisation, public reporting of performance data, adoption and use of health information technology, and other quality improvement initiatives. With the advent of artificial intelligence (AI), many believe there is potential for transformation of the health care industry through adoption of AI technologies.

Through a series of essays, this dissertation contributes to the literature on timeliness, equity, and efficiency in the emergency department. The first essay assesses recent trends in emergency department throughput in the United States. The second assesses current trends and sources of inequities in emergency department wait time

across racial and ethnic groups. In the third, a model aimed at improving efficiency in the emergency department is developed – an explainable machine learning model that leverages text data to predict patient disposition during triage is developed.

ESSAYS ON HEALTH CARE QUALITY: TIMELINESS, EQUITY, AND EFFICIENCY

by
Abubakar-Sadiq Bouda Abdulai

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Business Data Science

Martin Tuchman School of Management

August 2021

APPROVAL PAGE

ESSAYS ON HEALTH CARE QUALITY: TIMELINESS, EQUITY, AND EFFICIENCY

Abubakar-Sadiq Bouda Abdulai

Dr. Michael A. Ehrlich, Dissertation Advisor — Date
Associate Professor of Management, New Jersey Institute of Technology

---

Dr. Yi Chen, Committee Member — Date
Professor of Management, New Jersey Institute of Technology

Dr. Dantong Yu, Committee Member — Date
Associate Professor of Management, New Jersey Institute of Technology

---

Dr. Jorge Eduardo Fresneda Fernandez, Committee Member — Date
Assistant Professor of Management, New Jersey Institute of Technology

---

Dr. Zhi Wei, Committee Member — Date
Professor of Computer Science, New Jersey Institute of Technology

BIOGRAPHICAL SKETCH

Author: Abubakar-Sadiq Bouda Abdulai

Degree: Doctor of Philosophy

Date: August 2021

Undergraduate and Graduate Education:

- Doctor of Philosophy in Business Data Science,
  New Jersey Institute of Technology, Newark, NJ, 2021
- Master of Art in Economics,
  University of South Florida, Tampa, FL, 2016
- Master of Science in Mathematical Sciences,
  East Tennessee State University, Johnson City, TN, 2015
- Bachelor of Science in Actuarial Science,
  Kwame Nkrumah University of Science & Technology, Kumasi, Ghana, 2011

Major: Business Data Science

Presentations and Publications:

A. A. Sadiq Bouda and Ehrlich, M., "Impact of Affordable Care Act Medicaid Expansions on health-related quality of life of low-income adults." 10^{th} *Annual Conference of the American Society of Health Economists (ASHEcon). Virtual Conference. June 23, 2021.*

A. A. Sadiq Bouda, Mukhtar, F., and Ehrlich, M., "Decreased Racial and Ethnic Disparities in Emergency Department Wait Times in the United States." *Revise and Resubmit.*

A. A. Sadiq Bouda, Mukhtar, F., and Ehrlich, M., "United States' Performance on Emergency Department Throughput, 2006–2016." *Ann of Emer Med*, https://doi.org/10.1016/j.annemergmed.2021.01.009.

A. A. Sadiq Bouda and Ehrlich, M., "Explainable Artificial Intelligence (XAI): An Interpretable Model for Predicting Emergency Department Patient Disposition Using Verbatim Reason-for-Visit Narrative." *Submitted.*

To Mba Boknaba and Nma Adisa

My parents never went to school. But it was always clear from their actions and inactions that they believed it was the best for me. As a child, my mum would wake me up in the morning and prepare me for school, and my dad would hand me some money for breakfast. My mum was the economist. Occasionally, she would propose an increase in the amount due to "inflation". I remember vividly, in grade 3 (the year 1996), my father increased it from 50 to 60 cedis, thanks to my mum's advocacy.

They both couldn't read my performances as reported in my grade scripts but my brothers/cousins would always tell them I performed very well. They were always happy to hear that. Many would underestimate the conversations I had with my father since he died when I was only 14. But we did have some very intimate and life-changing conversations. One day, sitting in the front porch of our house with him and my cousin (Shaibu), he went on and on about life and living. That was before I was even a teenager. At a point, I responded in English, "there is no hurry in life". He did not understand, and asked my cousin what I said. My cousin quoted me in our local language. My father smiled, didn't look very surprised, as if to say I know my son will be fine. Then, he responded "yaa woto" (Indeed!).

It certainly took a lot of effort and time to complete this work. I wish he was here to witness this day. Nonetheless, I am grateful my mum is here with me today. She doesn't know what a PhD is, but she sure knows I have spent a heck of a long time in school that it must be something great. My parents gave me all I could have asked for: their support and my very supportive siblings. This work is dedicated to my mum and to the memory of my dad.

ACKNOWLEDGMENT

Praise and thanks to the Creator and Sustainer of the *Āllameen*, Allah (SWT), for His mercies. Indeed, His mercy encompasses everything (Qur'an 7:156).

I have benefited greatly from the advice and support of my advisor, Dr. Michael Ehrlich. I am grateful. I also appreciate the support and guidance of the members of my dissertation committee members, Dr. Yi Chen, Dr. Dantong Yu, Dr. Jorge Fresnada, and Dr. Zhi Wei. Thanks to you all.

Thanks to New Jersey Innovation Institute Healthcare Delivery Systems iLab for funding my research through a graduate research assistantship. Special thanks to Mr. Balavignesh Thirumalainambi for his constant support. Thanks to all the many hardworking staff of the Institute who supported me in numerous ways.

Thanks to Dr. Oya Tukel, for creating an enabling environment for students and research to thrive. Many thanks to the entire administrative staff of the Martin Tuchman School of Management, including Ms. Devoni Hinton, Ms. Lori Nanton, Ms. Virginia Dare, and Ms. Lilia Lozarito, for their support throughout the years. Thanks to lab mates Krystal Hunter, Milena Bajic, and Zhenrui Cao for their support. Their constructive criticism during our lab meetings enhanced my research.

Last, but certainly not least, thanks to my family for the unflinching support, especially, to Mr. Inusah Abdulai and Ms. Anne Kye-eebo. Thanks to all my friends for their support and prayers. Thanks to Ms. Jackie at the NJIT Campus Center for her support. Thanks to all who have made a contribution to my life in any way.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

According to the National Academy of Medicine (NAM), a quality health care system embodies six attributes: *timeliness, equity, safety, efficiency, effectiveness,* and *patient-centeredness* [1]. *Timeliness* is to avoid unnecessary delays in care delivery for patients and caregivers; *equity* is to ensure that the quality of care that patients receive does not vary based on personal characteristics such as race, ethnicity, socioeconomic status, geographic location, or other similar factors; *safety* is to ensure that the care that is intended to help patients does not harm them; *efficiency* is to avoid waste and optimize resource allocation to improve care delivery; *effective* care is one that relies on sound scientific knowledge to deliver the most benefit to the patient; and *patient-centeredness* is to put the patient first, respecting their choices, informing, and involving them and/or their designated family members in the care delivery process. To make progress towards achieving quality health care, efforts must be targeted towards improving every aspect of the health care system as related to the aforementioned attributes.

As recent as 2009, a review of data, the literature, and interview of emergency practitioners conducted by the Government Accountability Office (GAO) indicated that the state of timeliness in U.S. EDs was concerning. The report highlighted challenges in how long patients waited to be seen by a qualified emergency department (ED) practitioner - patients waited longer than recommended times, lengths of patient visits were longer, boarding of patients in ED hallways and other units for extended periods of time was prevalent, and crowding was pervasive and continued to occur over time [2].

While this was shown to be a general problem, further scrutiny from other researchers indicated that racial and ethnic minorities waited significantly longer than non-minorities to access care in United States (U.S.) EDs. In fact, racial and ethnic disparities are known to abound in many aspects of the US health care system [3]. In 2002, at the request of Congress, the Institute of Medicine (now National Academy of Medicine) published the findings and recommendations of a committee set up to investigate the extent of "unequal treatment" in the US healthcare system [4]. The resulting publication highlighted the nature and sources of racial and ethnic disparities. It also included recommendations for eliminating such disparities. Overall, the evidence suggests that minorities received inferior quality health care compared to non-minorities. The observed disparities were attributed to multiple factors including those related to patients and providers.

For many reasons, the US health system has been described as the most fragmented and inefficient system of care in the world [5]. For example, while the adoption and use of health information technology has been associated with improved care process management and efficiency, reduced cost of care, and improvements in quality [1], the US has only recently embraced the technological paradigm in its health care space. As at the end of 2010, only two-thirds of US EDs used basic IT systems and less than a third used advanced systems [6]. Inefficiencies in a fast-paced, stress-filled environment like the emergency department can have very devastating consequences. Especially, the complex relationship between the ED and other interconnected units of the health system complicates ED operations. The RAND Corporation previously estimated that nearly half of hospital admissions come from the ED [7]. The relationship between these two units (and others) implies that the repercussions of inefficiencies in one may affect the other. For example, boarding of patients in the ED is considered the primary contributor to crowding in (US) EDs [8]. Boarding may occur because there are no hospital beds available or patients cannot immediately access them due to lack of readiness on the part of hospital bed management. In the latter case, improved communication between ED staff and hospital bed management may significantly improve patient flow and thus reduce or eliminate boarding and therefore, crowding.

Patient-centeredness has recently gained impetus in the US health care system. This is a shift from the previously more traditional provider-centric approach to care delivery where caregivers do not actively engage patients and their family members in the care process. To practice patient-centeredness is to make the patient the "true north" of the care delivery process. The NAM defines patient-centered care as "health care that establishes a partnership among *practitioners, patients, and their families* (when appropriate) to ensure that decisions respect patients' wants, needs and preferences and that patients have the education and support they need to make decisions and participate in their own care" [9]. The role of the family in patient care cannot be overstated. It has been estimated that, unpaid family care provided by family care givers amounted to $450 billion in 2009 [10]. In general, evidence linking social ties and health has long been established by scientists. Individuals that are less socially connected are physically and psychologically less healthy and more likely to die when compared to individuals with strong social connections [11]. Thus, beyond direct medical care, engaging family members and encouraging them to continue to support patients both within and outside the care setting is necessary.

Urgent care settings such as the emergency and critical care units are high-stress, high-risk environments. As a result, there are many opportunities for error. Safety in US remains a national challenge. A survey of clininians identified problems in 4 systems critical to ED safety: physical environment, staffing, inpatient coordination, and information coordination and consultation [12]. Breakdowns in communication turn to mar teamwork and stifle the progress of care delivery and in some cases have contributed to adverse medical outcomes. The problem is compounded when these settings are crowded. Patients boarded in EDs become a distraction, and the pressure and stress resulting from caring for the ever-increasing number of patients increases the likelihood of error. ED staff have to make rapid clinical decisions but are significantly constrained by the amount of patient histories available to support decision-making which may lead to improper diagnosis - one of two of the most common errors in the ED.

Unlike the other five attributes, the U.S. is one of the global leaders of effective care delivery – one that is based on sound scientific knowledge. In particular, most emergency and trauma care settings (especially those at teaching hospitals) are generally able to assemble the most qualified and dedicated teams of experts, armed with the most sophisticated medical technology to provide care to patients. But this is not universal. These experts are known to be lacking in some locations (e.g., rural areas or specific geographic locations in urban areas) in which case the quality of care suffers. Besides, care provided in EDs may not be the most effective since doctors may not have access to the full patient medical history which implies that key information available to physicians is the patient's chief complaint offered at the point of contact, which may be inadequate.

1.2 Overview of the Research

As indicated above, a quality health care system provides timely, equitable, safe, efficient, effective, and patient-centered care. There remain unanswered questions on quality care delivery as related to these attributes. The goal of this research is to contribute to the literature on health care quality through a series of essays that focus on three of the six mutually non-exclusive attributes described above.

This research starts off with an evaluation of the current state of timeliness in US EDs using national data. To provide new tools for improving timeliness, it is important to evaluate the current state of the challenge.

Inequities in ED patient wait time times are also assessed. Following enormous documentation of long waits and inequities within, recent federal-, state-, and hospital-level initiatives were targeted towards improving timeliness and the quality of care in general. Whether these recent reforms have had any impact on previously observed inequities in patient wait time remains unknown.

An interpretable deep learning model is developed to predict patient disposition using qualitative data collected at triage. The goal is to replace a physician-in-triage with an automated model. To improve patient flow and overall efficiency, a common practice by EDs is to place a physician in triage to undertake a quick assessment of a patient and make a tentative decision about the patient's most likely disposition. This is aimed at improving communication between the admissions and bed management teams to ensure easy transfer of admitted patients to inpatient wards. This process can be automated with an algorithm. In this section, such model is developed. This model uses data on all patients from heterogeneous EDs, and unlike any other, incorporates verbatim reason for visit narrative text in the prediction task. For patients that are admitted, the explainable model will further identify which key words in a given sequence of text explain the variability in patient disposition.

In summary, this dissertation will contribute to the literature on quality of care in a very important urgent care setting, the emergency department. The first assesses recent trends in timelines and the rate of patient flow in US emergency departments. The second assesses the current state and source of inequities of ED patient wait time. The third contributes to improvement in timeliness and efficiency in the ED via predictive model development.

1.3 Organization of the Dissertation

The next three chapters are essays on three mutually non-exclusive attributes of quality. In Chapter 2, the current state of timeliness in US EDs is assessed. The literature on the state of ED timeliness nationally, is reviewed and new results on the current situation are provided and discussed. In Chapter 3, the problem of prolonged waiting to see the emergency physician in the US is introduced. The literature on racial and ethnic disparities in ED wait times is discussed. Current trends and sources of the disparities are investigated and results presented and discussed. The potential impact of recently proposed throughput strategies targeted towards improving timeliness in US EDs is also assessed. Chapter 4 presents a proposed interpretable model for predicting patient disposition at triage using verbatim patient reason for visit narrative text.

CHAPTER 2

AN ASSESSMENT OF EMERGENCY DEPARTMENT THROUGHPUT

"Defer no time, delays have dangerous ends" – Henry VI, Shakespeare

2.1 Introduction

A 2006 report from the National Academy of Medicine (NAM) on crowding in U.S. hospital-based emergency departments (EDs) described the situation as a national epidemic [13]. The concerns highlighted included EDs operating at or above capacity, long waiting times to access care, persistent patient "boarding" (i.e., patients kept waiting in ED hallways or other space outside the ED on admission to hospital), and ambulance diversion. These phenomena have been associated with patient dissatisfaction [14, 15, 16, 17], loss of revenue [18, 19], increased medical costs [20, 21, 22] and worse health outcomes [22, 23, 24, 25, 26, 27]. The IOM report thus signaled a clarion call to remedy the overburdened status of EDs, which was nearing a "breaking point." Yet, another report from the Government Accountability Office (GAO) published three years after the IOM report, indicated crowding continued to occur in U.S. EDs [2].

Since the release of these reports, several national initiatives have been implemented to improve ED timeliness. For instance, a series of ED performance and benchmarking summits were held in 2006, 2010, and 2014, where leaders in emergency medicine deliberated ED process measurement and performance improvement [28, 29, 30]. In 2008, a task force set up by the American College of Emergency Physicians proposed a series of "high-impact solutions" for reducing crowding and improving timeliness in U.S. EDs [8]. In October 2013, the Centers for Medicare and Medicaid Services (CMS) included measures of ED timeliness in the

Hospital Inpatient Quality Reporting Program intended to enhance ED throughput. Under this program, hospitals that report these measures to CMS earn a higher annual update to their payment rates, and those that fail to report them receive a reduction in payment [31]. The reported measures are also publicly reported on the hospital compare website [32].

The most recent annual ED summary tables published by the Center for Disease Control and Prevention (CDC) suggest that ED presentations have increased in recent times [33]. Also, the recent expansions of Medicaid following the implementation of the Affordable Care Act (ACA) in 2014 has been associated with increased use of ED services, with concerns about potential ED crowding [34, 35]. While a few previous studies have evaluated national trends in ED throughput in the US, these studies span periods that pre-date the IOM or GAO reports and assessed only one or two throughput measures [36, 37, 38, 39, 40, 41, 42]. Thus current national trends in ED throughput remain unknown.

In this study, trends in five measures of ED timeliness are assessed: (1) median time from door to diagnostic evaluation by a qualified ED practitioner (i.e., wait time); (2) percent leaving without being seen (LWBS); (3) median time from arrival to departure for discharged patients; (4) median time from arrival to departure for admitted patients; and (5) median time from admission decision to ED departure for admitted patients (i.e., boarding time). In addition to these, trends in the proportion seen within recommended waiting times [43, 44], and the proportion seen within four, six, and eight hours as practiced in other developed countries [45, 46, 47, 48]. While an assessment of the proportion of patients dispositioned within such target hours is a measure of throughput in itself, it is also interesting to see how the U.S. performs in the absence of these defined time benchmarks.

2.2 Literature Review

A few studies have evaluated ED timeliness or throughput on a national scale. Horwitz and Bradley assessed the rate at which patients were seen by an ED physician within recommended times [39]. They found that between 1997 and 2006, the proportion of patients seen within the target wait times declined by a mean of 0.8% anually. They found that the most affected was the sickest group of patients who should be seen within 14 minutes. During the study period, the proportion of the sickest patients seen within the said recommended time decreased from 59.2% to 48.0%. They concluded that, the percent of patients seen within recommended times was "steadily declining and ... at its lowest point in at least 10 years."

Pitts et al. evaluated national trends in ED crowding and its potential causes by analyzing occupancy rates from 2001 to 2008 [36]. They found that during that period, the number of ED presentations increased at a rate 60% faster than the U.S. population growth rate during that period. Average occupancy rates were even much higher. They concluded that crowding in U.S. EDs was getting worse. That, while sociodemographic dynamics over time accounted for some of the observed trends, most of it resulted from practice intensity (which includes the use of intravenous fluids, blood test, computed tomography (CT) scans, magnetic resonance imaging (MRI), and ultrasonographic examinations).

Pitts et al. also assessed trends in boarding duration for individual patient visits from 2007 to 2010 using national data [37]. They found that national median boarding time exceeded an hour and nearly a third of admitted visits were boarded for more than 2 hours. Even among EDs that indicated that they did not board patients for more than 2 hours, the data indicated that 21% of admissions at these EDs did stay in the ED over 2 hours.

2.3 Methods

2.3.1 Study Design and Data

This was a retrospective cross-sectional study of patient encounters at hospital-based EDs in the U.S. from 2006 to 2016. Data were from the ED subsample of the National Hospital Ambulatory Medical Care Survey (NHAMCS). The NHAMCS is a survey conducted by the National Center for Health Statistics (NCHS), a division of the Centers for Disease Control and Prevention (CDC).

The survey collects data on emergency care utilization across all 50 states and the District of Columbia. It contains survey design variables that allow for estimates at the national and sub-divisional levels using survey weights provided by NCHS. The publicly available files, which have been de-identified and masked to reduce disclosure risk for patients and EDs were used in this study.

2.3.2 Data Collection and Processing

In NHAMCS, a four-stage sampling procedure is used to collect data. In the first stage, NCHS samples 112 probability sampling units (PSUs) (which are counties, county equivalents, towns/townships, or metropolitan areas) from a sample of 1,900 used in the National Health Interview Survey. From the 112 PSUs, up to 600 short-stay (average of 30 days) hospitals are sampled. In the third stage, emergency service areas (ESAs) in the sampled hospitals are identified as the ED subsample. Finally, patient encounters are sampled from these EDs.

During a randomly assigned 4-week reporting period, ED staff are requested to keep a list of all patient encounters, from which a systematic random sample of 100 patients is taken. Data are then electronically (since 2012) abstracted from patient records. Data collected include patient demographics including age, sex, race/ethnicity, payment source; clinical details including reason for visit, physician diagnosis, prescribed medications; and other patient and provider information. More information about NHAMCS survey is available on the CDC's website [49]

2.3.3 Outcome Variables

The primary outcome variables included wait time, length of visit (LOV), boarding time, and the proportion leaving (before or after triage) without being seen (LWBS) by a qualified practitioner. Two other computed outcomes were the proportion seen within recommended waiting times, which was computed from the wait time measure, and the proportion seen within 4, 6, and 8 hours, which was computed from LOV. Wait time is the time (in minutes) from arrival to meeting with a qualified ED practitioner. LOV is the time (in minutes) from arrival to disposition (time admission decision is made or physically leaving ED on discharge). Boarding time is the time (in minutes) from bed request for hospital admission or transfer to the time the patient leaves the ED or observation unit. The boarding time outcome was first recorded in 2009. Due to data quality issues, this outcome was not published in 2012 and 2016. LOV data was also not published in 2016 due to quality issues. As a result, the analysis of wait time was from 2006 to 2016, that of LOV from 2006 to 2015, and that of boarding time from 2009 to 2015 (excluding 2012). None of the outcome variables were imputed since the response rates for these variables generally exceeded 80% which is very high for a survey of this nature [50]. LWBS had no missing data.

2.3.4 Independent/Control Variables

The independent variables were patient age, sex, race/ethnicity, source of payment, U.S. region, metropolitan status, work shift, season, and patient disposition. Dummy variables were created for weekday, resident intern involvement in care, diagnostic testing, and psychiatric reason for visit (psychological and mental disorder). Some of the survey items including patient age, sex, race, and ethnicity with missing data were imputed by the NCHS before public release. In years prior to 2013, triage level was also imputed. The independent variables missing data were also not imputed.

2.3.5 Primary Data Analysis

Data were characterized using standard descriptive statistics. The annual distributions of wait time, length of visit, and boarding time variables were graphed using box-and-whisker plots. Since the data were skewed, the lower fence/whisker of the boxplot was defined as maximum (0, $Q_1 - 3{*}SIQR_L$) and the upper fence/whisker was defined as $Q_3 + 3{*}SIQR_U$ [51]. $SIQR_L$, the lower semi-interquartile range is $Q_2 - Q_1$, and $SIQR_U$, the upper semi-interquartile range is $Q_3 - Q_2$. Q_1, Q_2, and Q_3 are the first, second, and third quartiles. This approach is conservative and data points outside these fences/whiskers must be far outside [52].

The weighted total number of visits, medians, and interquartile ranges of wait time and length of visit for the first and last years of the study period were estimated. This was stratified by patient age (0−17, 18−24, 25−44, 45−64, and 65+ years), sex, race/ethnicity (White, non-Hispanic; Black, non-Hispanic; Hispanic; Other, non-Hispanic), payment (private insurance, Medicare, Medicaid/CHIP, self-pay, other), Region (Northeast, Midwest, South, West), metropolitan statistical area (yes, no), patient disposition (admitted, discharged), triage level (emergent, urgent, semi-urgent, non-urgent, no triage, unknown) and the primary reason for visit (psychiatric, non-psychiatric). Others were work shift (7:00-15:00; 15:00-23:00; 23:00:7:00), weekday (yes, no), season (winter, spring, summer, autumn), and the use of diagnostic services during a visit (yes, no).

To assess changes in median wait time and median length of visit over time, unadjusted regression models were fit through the median of the combined data from 2006 to 2016 for these outcomes. The outcomes were wait time and length of visit, and the independent variable was year of visit. The significance in trend in the median was assessed via the significance of the coefficient for the year variable in the model. The trend in proportion leaving without being seen (LWBS) was also analyzed by fitting an unadjusted logistic regression model to the combined data from 2006 to 2016. The outcome was (binary) LWBS, and year of visit was the independent variable. The significance in the trend was assessed via the significance of the coefficient for the year variable in the model.

The proportion of visits seen within recommended waiting times was also computed. The odds of a patient being seen within those times were also computed by fitting logistic regression models to the whole as well as each of the four triage subgroups (emergent, urgent, semi-urgent, and non-urgent). Patients presenting to the ED were assigned one of five severity scores (1 to 5) based on an assessment of the degree of urgency required to attend to the patient. A score of 1 is defined as "immediate", where any delays may result in death (see patient in 0 minutes). A score of 2 is defined as "emergent", where delays would likely result in deterioration of the patient's condition (see the patient within 14 minutes). A score of 3 is defined as "urgent", where treatment should be provided as soon as possible (see the patient within 60 minutes). A score of 4 is defined as "semi-urgent", where illness or injury requires treatment within 120 minutes. A score of 5 is defined as "non-urgent", where a delay of up to 24 hours would not make a significant difference in the patient's condition (see within 24 hours). For the approximately 10% of records where a $1-3$ or $1-4$ triage scale was used, the NCHS rescaled these to $2-4$ and $2-5$ scales, respectively. Triage levels 1 and 2 merged and labeled "emergent" (see the patient within 14 minutes) to facilitate analysis and comparison with previous results. In the logistic model, the outcome was 1 if the patient was seen within the recommended waiting time for the respective acuity level, and 0 otherwise. The independent variable was year of visit (2006 to 2016).

The proportion of visits dispositioned within 4, 6, and 8 hours of arrival to the ED and trends within were estimated. This analysis was stratified by admission disposition and triage category. Changes in trends were assessed within each triage category by fitting unadjusted and adjusted logistic regression models to the 2006−2015 combined data. The outcome was 1 if the length of visit during a visit was within the specified hour (4, 6, or 8), and 0 otherwise. In the adjusted models, the covariates were the year of visit, patient age, sex, race, source of payment, U.S. region, metropolitan status, weekday, work shift, season, and whether or not diagnostic services were provided during the visit. There was a separate model for each acuity level, and so patient acuity was not adjusted for. Finally, the absolute percent changes in the total number of visits between 2006 and 2016 were estimated.

In all analyses, outliers were not excluded from the analysis because the median is resistant to outliers. A second wave of analysis for the primary outcomes was limited data to the 99^{th} percentile of annual data to minimize the influence of outliers. It was found that the results were very similar. The multistage probability design of NHAMCS was accounted for as suggested by NCHS using the Survey package (version 3.22) [53] in the R software (version 2.11) [54]. Where estimates had standard errors greater than or equal to 30% of point estimates or the analysis involved a subsample with less than 30 observations, such estimates were ignored due to the large uncertainty associated with such estimates. This is in accordance with the suggestions of the NCHS.

2.4 Results

Table 2.1. Sample Characteristics by Patient and Provider Subgroups, 2006-2016

	Full NHAMCS sample		**Length of visit**	**Wait time**
Characteristic	**2006-2015**	**2006-2016**	**2006-2015**	**2006-2016**
All, N ()	305,570 (100)	325,037 (100)	288,662 (100)	272,731 (100)
Age group, n(%)				
0−17	69,152 (22.6)	73,377 (22.6)	65,332 (22.6)	61,225 (22.4)
18−24	36,696 (12.0)	38,803 (11.9)	34,581 (12.0)	32,062 (11.8)
25−44	87,227 (28.5)	92,686 (28.5)	82,471 (28.6)	77,273 (28.3)
45−64	66,480 (21.8)	71,115 (21.9)	62,615 (21.7)	59,896 (22.0)
65+	46,015 (15.1)	49,056 (15.1)	43,663 (15.1)	42,275 (15.5)
Sex, n(%)				
Female	166,212 (54.4)	176,787 (54.4)	157,212 (54.5)	148,539 (54.5)
Male	139,358 (45.6)	148,250 (45.6)	131,450 (45.5)	124,192 (45.5)
Race/Ethnicity, n(%)				
White	178,509 (58.4)	189,999 (58.5)	169,567 (59.7)	159,489 (58.5)
Black	68,778 (22.5)	73,048 (22.5)	64,820 (22.5)	61,719 (22.6)
Hispanic	45,531 (14.9)	48,409 (14.9)	42,247 (14.4)	40,196 (14.7)
Other	12,752 (4.2)	13,581 (4.2)	12,028 (4.2)	11,327 (4.2)
Payment Source, n(%)				
Private	92,592 (30.3)	97,763 (30.1)	87,893 (30.4)	82,379 (30.2)
Medicare	51,681 (16.9)	55,261 (17.0)	49,145 (17.0)	47,770 (17.5)
Medicaid/CHIP	85,254 (27.9)	92,230 (28.4)	80,449 (27.9)	78,253 (28.7)
Self-Pay	41,075 (13.4)	42,602 (13.1)	39,066 (13.5)	35,280 (12.9)
Other	34,968 (11.4)	37,181 (11.4)	32,109 (11.1)	29,049 (10.7)
Region, n(%)				
Northeast	69,894 (22.9)	72,907 (22.4)	65,156 (22.6)	59,378 (21.8)
Midwest	69,273 (22.7)	73,770 (22.7)	66,690 (23.1)	62,737 (23.0)
South	106,456 (34.8)	113,133 (34.8)	101,374 (35.1)	97,800 (35.9)
West	59,947 (19.6)	65,227 (20.1)	55,442 (19.2)	52,816 (19.4)
Metropolis†, n(%)				
Yes	267,557 (87.6)	284,265 (87.5)	199,388 (77.6)	213,486 (78.3)
No	38,013 (12.4)	40,772 (12.5)	36,780 (12.7)	34,662 (12.7)
Disposition, n(%)				
Admitted	36,886 (12.1)	38,502 (11.8)	34,767 (12.0)	33,397 (12.2)
Discharged	268,684 (87.9)	286,535 (88.2)	253,895 (88.0)	239,334 (87.8)
Triage level				
Emergent	36,798 (12.0)	38,432 (11.8)	35,078 (12.2)	33,864 (12.4)
Urgent	115,199 (37.7)	121,725 (37.4)	110,351 (38.2)	106,371 (39.0)
Semi-urgent	81,293 (26.6)	86,088 (26.5)	78,048 (27.0)	73,952 (27.1)
Non-urgent	22,931 (7.5)	23,790 (7.3)	21,796 (7.6)	19,518 (7.2)
No triage	14,694 (4.8)	15,828 (4.9)	12,788 (4.4)	11,446 (4.2)
Unknown/Missing	34,655 (11.3)	39,174 (12.1)	30,601 (10.6)	27,580 (10.1)
Reason for visit, n(%)				
Psychiatric	10,379 (3.4)	11,078 (3.4)	9588 (3.4)	9,210 (3.5)
Non-psychiatric	295,191 (96.6)	313,959 (96.6)	279,074 (96.6)	263,521 (96.5)

Notes: Abbreviations: CHIP, Children's Health Insurance Program. †Metropolitan status was not publicly reported in 2012. Definitions: Emergent: $<$ 15 minutes; urgent: $\leq$ 60 minutes; semi-urgent: $\leq$ 2 hours; non-urgent: $\leq$ 24 hours.

2.4.1 Characteristics of Study Subjects

The full NHAMCS sample for years 2006 to 2016 was 325,037. Of those, wait time was documented for 272,731 (83.9% response rate). Since LOV analysis was limited to the period 2006 to 2015, the full NHAMCS sample size of 305,570 for this period is also reported. Patient length of visit was recorded for 288,662 encounters (94.5% response rate) (Table 2.1). As can be seen, the sample characteristics of the analytical samples (for wait time and length of visit) are similar across patient and provider subgroups.

2.4.2 Number of Visits

From 2006 to 2016, the total number of visits to EDs increased from 119.2 million to 145.6 million, an absolute change of 29.4 million visits (Table 2.2). This change represents an increase of 22.1%, significantly higher than the U.S. population growth of 8.5% during that period [55]. Across age groups, visits increased the most among those aged 45 years or more, while those aged 18−24 years had the least increase in visits. Across racial and ethnic groups, Whites had an increase of 16 million visits, the most in the group. "Other" race had less than a 0.1 million increase in visits. There was a 3.4 million decline in number of visits where private insurance was presented as the source of payment. Those presenting with Medicaid on the other hand, increased by nearly 19.1 million. Visits where patients self-paid declined by 7.6 million during the period. Those presenting with Medicare coverage increased by 9.1 million. Regionally, the West experienced the most increase in presentations, which stood at 16.1 million. The Northeast experienced the least increase, at 1.8 million. Visits to non-metropolitan areas increased by 9.9 million compared to 16.5 million in metropolitan areas.

Table 2.2. Trends in Total Number of Visits, Millions (Standard Error), 2006-2016

Characteristic	**2006**	**2016**	**Absolute Change**
Sample, N	35,849	19,467	
All	119.2 (5.8)	145.6 (8.9)	29.4
Age group			
0−17	26.3 (1.4)	32.1 (3.0)	5.8
18−24	15.1 (0.9)	16.0 (1.0)	0.9
25−44	35.0 (1.9)	40.0 (2.6)	5.0
45−64	25.5 (1.4)	34.4 (2.2)	8.9
65+	17.3 (0.8)	23.1 (1.8)	5.8
Sex			
Female	65.0 (3.3)	79.6 (4.9)	14.6
Male	54.2 (2.5)	66.0 (4.1)	11.8
Race/Ethnicity			
White	71.9 (3.9)	87.9 (6.0)	16.0
Black	28.1 (2.8)	30.7 (3.1)	2.6
Hispanic	14.8 (1.5)	22.4 (2.5)	7.6
Other	4.4 (0.5)	4.5 (0.7)	0.1
Source of Payment			
Private	40.0 (2.1)	36.6 (2.6)	−3.4
Medicare	16.8 (0.9)	25.9 (1.8)	9.1
Medicaid/CHIP	30.3 (1.8)	49.4 (3.3)	19.1
Self-Pay	19.3 (1.3)	11.7 (1.3)	−7.6
Other	12.8 (1.2)	22.0 (4.9)	9.2
Region			
Northeast	22.7 (1.7)	24.5 (4.1)	1.8
Midwest	25.7 (3.6)	31.4 (3.5)	5.7
South	50.6 (4.1)	53.5 (5.3)	2.9
West	20.1 (1.4)	36.2 (4.8)	16.1
Metropolis*			
Yes	100.7 (7.9)	117.2 (10.6)	16.5
No	18.5 (4.8)	28.4 (7.2)	9.9
Disposition			
Admitted	15.2 (1.0)	12.6 (1.3)	−2.6
Discharged	104.0 (5.1)	132.9 (8.2)	28.9
Triage level			
Emergent	20.3 (1.5)	23.1 (2.3)	2.8
Urgent	43.6 (3.3)	47.2 (3.5)	3.6
Semi-urgent	26.2 (2.2)	35.7 (2.9)	9.5
Non-urgent	14.5 (1.6)	6.3 (0.8)	−8.2
Reason for visit			
Psychiatric	3.1 (0.3)	4.8 (0.4)	1.7
Non-psychiatric	116.1 (5.6)	140.8 (8.7)	24.1

Patient encounters that resulted in admission to the hospital ward declined by 2.6 million, while those discharged increased by 28.9 million during the period. The group triaged as semi-urgent had the most change in visits, an increase of 9.5 million. The non-urgent group had an 8.2 million decline in visits. The most acute groups, emergent and urgent, had increases of 2.8 million and 3.6 million visits, respectively. Visits where psychological and mental disorders were identified as the primary reason for visiting increased by 1.7 million while non-psychiatric visits increased by 24.7 million. Other variables by which the number of visits was characterized are work shift (morning, afternoon, evening), weekday (yes/no), season, and whether or not physician ordered diagnostic services. These results are available in Table A.1 in Appendix A.

2.4.3 General Distribution of Waiting Time

The yearly distributions of wait time are shown in Figure 2.1. Wait times had an upward trend from 2006 to 2009, during which the median was above 30 minutes. The period after that had a downward trend, during which the median was below 30 minutes. The largest non-outlier (upper fence/whisker value of boxplot) wait time across all years was 193 minutes (or 3.2 hours) observed in 2009.

Estimates of the median and interquartile range for 2006 and 2016, as well as the year coefficient (with confidence intervals) for trend (from 2006 to 2016) assessment, are shown in Table 2.3. Overall, median wait time decreased at an annual rate of 1.9 minutes (95% CI: 2.3, 1.5), from 31 minutes (IQR, 14 to 67) to 17 minutes (IQR, 6 to 45). Declines in the median were fairly similar across age and sex groups. All racial/ethnic groups experienced significant declines in median wait times. As at 2016, no major differences existed in the medians and interquartile ranges of wait time across all racial/ethnic groups. Blacks who initially had the longest median

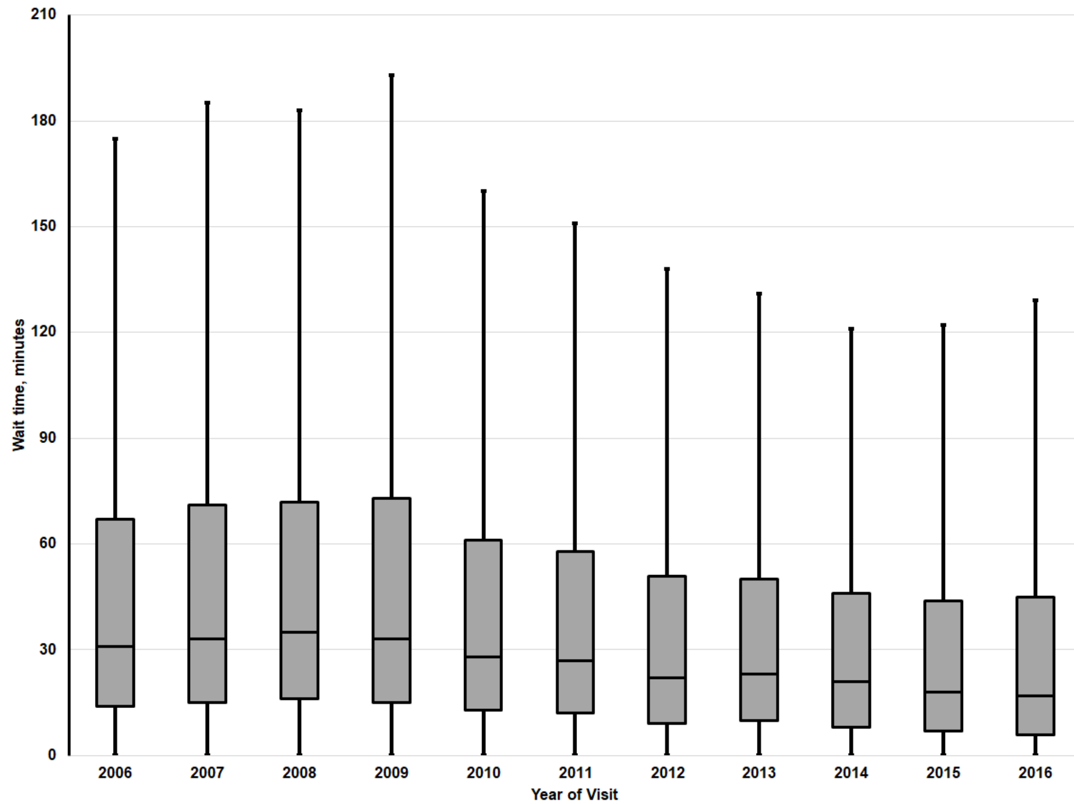

Figure 2.1. Yearly distribution of wait time (in minutes) for all visits, 2006−2016. **Note:** NHAMCS data were weighted to be nationally representative.

waiting time, experienced an annual decline 2.8 minutes over time (39 minutes [IQR: 18 to 83] to 18 minutes [IQR: 5 to 49]).

There were also declines in median wait time across all the other subgroups. Declines across modes of payment ranged from 1.4 minutes (95% CI: 1.0 to 1.8) for Medicare holders to 2.6 minutes (95% CI: 1.9 to 3.2) for "other" source of payment. The declines did not vary much by region and urban status, and which were quite similar to the national trend. Having stratified by patient disposition, median wait time for the admitted decreased from 25 minutes (IQR: 10 to 57) to 18 minutes (IQR: 6 to 44), and that of the discharged decreased from 33 minutes (IQR: 15-69) to 17 minutes (IQR: 6 to 45). Across acuity levels, median wait time for the emergent group increased from 15 minutes (IQR: 6 to 37) to 18 minutes (IQR: 6 to 43), albeit insignificant (annual change of 0.1 minutes [95% CI: −0.4 to 0.6]). But the remaining

Table 2.3. Trends in Median (Interquartile Range) Wait Time in Minutes, 2006-2015

Characteristic	2006	2015	Annual Change (95% CI)
Sample, n	28,391	16,310	
All	31 (14 to 67)	17 (6 to 45)	−1.9 (−2.3 to −1.5)
Age group			
0−17	33 (15 to 66)	18 (6 to 41)	−2.0 (−2.4 to −1.6)
18−24	35 (15 to 75)	17 (5 to 42)	−2.2 (−2.8 to −1.6)
25−44	33 (15 to 71)	18 (6 to 49)	−2.0 (−2.5 to −1.5)
45−64	31 (14 to 66)	18 (6 to 47)	−1.8 (−2.2 to −1.3)
65+	25 (10 to 59)	15 (5 to 41)	−1.3 (−1.8 to −0.9)
Sex			
Female	33 (15 to 70)	17 (6 to 45)	−2.0 (−2.5 to −1.5)
Male	30 (14 to 65)	17 (5 to 44)	−1.7 (−2.1 to −1.2)
Race/Ethnicity			
White	29 (12 to 60)	17 (6 to 42)	−1.5 (−1.9 to −1.1)
Black	39 (18 to 83)	18 (5 to 49)	−2.8 (−3.5 to −2.1)
Hispanic	36 (17 to 76)	17 (6 to 47)	−2.2 (−2.8 to −1.5)
Other	29 (12 to 60)	13 (5 to 48)	−2.0 (−2.7 to −1.3)
Source of Payment			
Private	32 (14 to 65)	17 (5 to 44)	−1.9 (−2.3 to −1.4)
Medicare	28 (11 to 60)	16 (5 to 45)	−1.4 (−1.8 to −1.0)
Medicaid/CHIP	33 (15 to 68)	20 (7 to 49)	−1.7 (−2.3 to −1.1)
Self-Pay	35 (16 to 76)	20 (7 to 52)	−2.0 (−2.6 to −1.4)
Other	30 (13 to 67)	13 (5 to 32)	−2.6 (−3.2 to −1.9)
Region			
Northeast	34 (16 to 69)	24 (9 to 48)	−1.8 (−2.6 to −1.0)
Midwest	29 (13 to 58)	15 (4 to 39)	−1.6 (−2.5 to −0.8)
South	34 (15 to 75)	18 (5 to 46)	−2.0 (−2.7 to −1.3)
West	27 (11 to 57)	15 (5 to 44)	−1.5 (−2.3 to −0.7)
Metropolis*			
Yes	35 (15 to 74)	20 (7 to 50)	−2.0 (−2.5 to −1.5)
No	20 (10 to 40)	9 (2 to 25)	−1.2 (−2.1 to −0.3)
Disposition			
Admitted	25 (10 to 57)	18 (6 to 44)	−1.0 (−1.5 to −0.5)
Discharged	33 (15 to 69)	17 (6 to 45)	−2.0 (−2.5 to −1.5)
Triage level			
Emergent	15 (6 to 37)	18 (6 to 43)	0.1 (−0.4 to 0.6)
Urgent	31 (15 to 59)	19 (6 to 49)	−1.7 (−2.1 to −1.2)
Semi-urgent	45 (21 to 89)	19 (7 to 49)	−2.9 (−3.5 to −2.2)
Non-urgent	45 (21 to 99)	18 (6 to 39)	−3.1 (−4.3 to −2.0)
Reason for visit			
Psychiatric	36 (13 to 88)	16 (5 to 37)	−2.0 (−2.8 to −1.2)
Non-psychiatric	31 (14 to 67)	17 (6 to 45)	−1.8 (−2.2 to −1.4)

acuity groups experienced statistically significant declines in median wait times. The urgent triage group experienced an annual decline 1.7 minutes (95% CI: 1.2 to 2.1) in the median, decreasing from 31 minutes (IQR: 15 to 59) to 19 minutes (IQR: 6 to 49). The semi-urgent group experienced an annual decline of 2.9 minutes (95% CI: 2.2 to 3.5) in the median, decreasing from 45 minutes (IQR: 21 to 89) to 19 minutes (IQR: 7 to 49). And the non-urgent group experienced an annual decline of 3.1 minutes (95% CI: 2.0 to 4.3) in the median, decreasing from 45 minutes (IQR: 21 to 99) to 18 minutes (IQR: 6 to 39). This suggests that over time, previously observed differences in median wait times across triage levels disappeared. Among those visiting with psychiatric chief complaints, median wait times declined at an annual rate of 2 minutes (95% CI: 1.2 to 2.8), not very different from the 1.8 minutes (95% CI: 1.4 to 2.2) decline experienced by those with non-psychiatric-related chief complaints. Other variables by which the patient wait time was characterized are work shift, weekday (yes/no), season, and whether a physician ordered diagnostic services. These results are available in Table A.2 in Appendix A.

2.4.4 Length of Visit

Length of visit for the Admitted. The distribution of LOV for the admitted remained fairly unchanged across all years from 2006 and 2015 (Figure 2.2). Across all years, the largest non-outlier LOV for the admitted was approximately 987 minutes or 16.5 hours (in 2012). There was no statistically significant change in median LOV for the admitted (annual change of 0.8 minutes [95% CI: -1.6 to 3.2]) (Table 2.3).

Length of visit for the Discharged. The distribution of LOV for the discharged also remained fairly unchanged across all years from 2006 and 2015 (Figure 2.3). Across all years, the largest non-outlier LOV for the discharged was approximately 556 minutes or 9.3 hours (in 2012). There was no statistically significant change in

Table 2.4. Trends in Median (Interquartile Range) Length of Visits in Minutes, 2006-2015

Characteristic	**2006**	**2015**	**Annual Change (95% CI)**
Sample, n	33,342	19,562	
All	149 (85 to 244)	154 (91 to 249)	−0.2 (−1.4 to 0.9)
Age group			
0−17	113 (69 to 184)	115 (70 to 182)	−0.5 (−1.5 to 0.5)
18−24	135 (81 to 224)	131 (78 to 223)	−0.8 (−2.4 to 0.8)
25−44	143 (80 to 241)	152 (90 to 246)	0.0 (−1.5 to 1.5)
45−64	174 (102 to 280)	183 (111 to 289)	−0.3 (−2.1 to 1.5)
65+	195 (125 to 300)	208 (135 to 309)	0.0 (−1.9 to 1.9)
Sex			
Female	155 (88 to 248)	159 (94 to 251)	−0.5 (−1.9 to 0.9)
Male	141 (81 to 236)	149 (87 to 246)	0.1 (−1.1 to 1.4)
Race/Ethnicity			
White	160 (91 to 265)	151 (92 to 248)	0.5 (−0.9 to 1.9)
Black	145 (83 to 237)	155 (91 to 248)	−2.3 (−4.0 to −0.6)
Hispanic	155 (90 to 259)	161 (92 to 256)	0.0 (−2.0 to 2.0)
Other	152 (82 to 248)	161 (87 to 274)	−2.5 (−6.3 to 1.3)
Source of Payment			
Private	142 (83 to 232)	156 (92 to 246)	0.7 (−1.0 to 2.4))
Medicare	193 (120 to 296)	198 (126 to 306)	−0.3 (−2.4 to 1.8))
Medicaid/CHIP	138 (79 to 228)	139 (82 to 229)	−0.7 (−2.2 to 0.9))
Self-Pay	145 (83 to 242)	133 (78 to 225)	−2.0 (−3.8 to −0.2))
Other	140 (78 to 244)	147 (91 to 235)	0.2 (−1.7 to 2.0))
Region			
Northeast	151 (85 to 246)	181 (104 to 296)	2.0 (0.1 to 3.9)
Midwest	134 (75 to 228)	161 (96 to 251)	0.3 (−2.7 to 3.3)
South	155 (90 to 249)	135 (79 to 223)	−1.3 (−4.2 to 1.5)
West	151 (84 to 245)	165 (97 to 256)	0.2 (−2.8 to 3.1)
Metropolis*			
Yes	158 (90 to 258)	163 (97 to 261)	−0.7 (−2.1 to 0.8)
No	108 (63 to 172)	111 (64 to 180)	1.3 (−0.6 to 3.1)
Disposition			
Admitted	260 (169 to 385)	278 (192 to 409)	0.8 (−1.6 to 3.2)
Discharged	136 (80 to 221)	146 (87 to 233)	0.3 (−0.9 to 1.6)
Triage level			
Emergent	170 (100 to 272)	225 (151 to 335)	6.4 (3.9 to 8.9)
Urgent	160 (93 to 255)	195 (126 to 296)	2.7 (1.1 to 4.2)
Semi-urgent	137 (79 to 223)	115 (71 to 186)	−3.0 (−4.3 to −1.7)
Non-urgent	124 (70 to 207)	89 (55 to 153)	−4.8 (−6.8 to −2.8)
Reason for visit			
Psychiatric	242 (139 to 414)	258 (150 to 402)	1.8 (−2.2 to 5.9)
Non-psychiatric	147 (84 to 240)	152 (90 to 244)	−0.3 (−1.6 to 1.1)

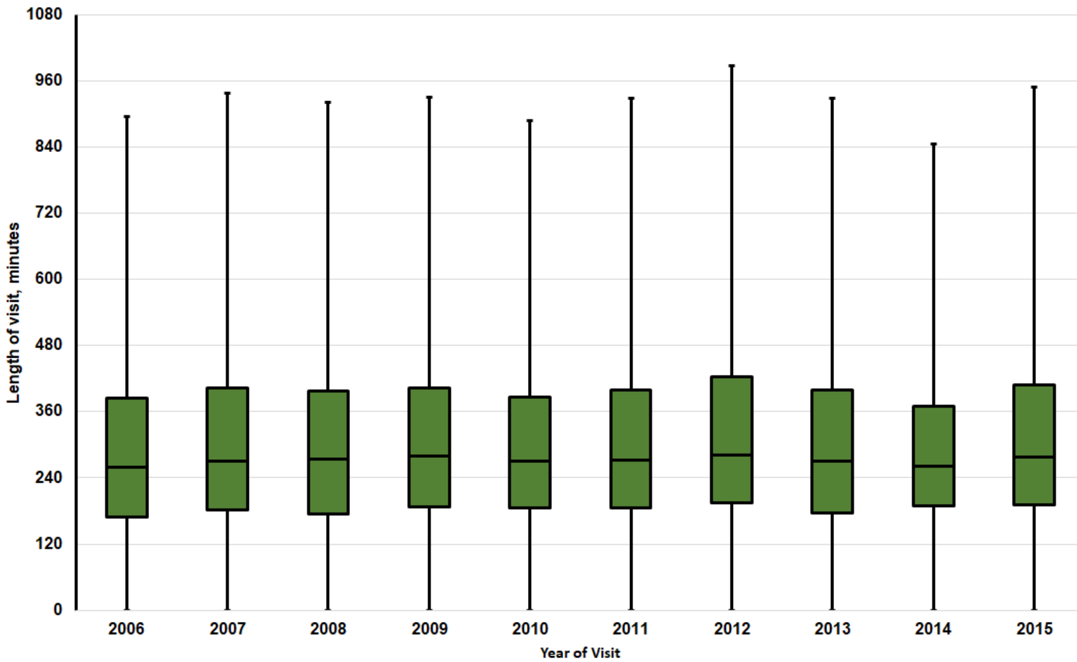

Figure 2.2. Yearly distribution of length of visit (in minutes) for admitted patients, 2006−2015. **Note:** NHAMCS data were weighted to be nationally representative.

median LOV for the admitted (annual change of 0.3 minutes [95% CI: -0.6 to 3.1]) (Table 2.3).

General Trends in Length of Visit. Overall, there was no statistically significant change in median LOV between 2006 and 2015 (annual change of -0.2 (95% CI: -1.4 to 0.9). Median LOV for all patients was 149 minutes (IQR: 85 to 244) in 2006 and 154 minutes (IQR: 91 to 249) in 2015, a difference of 5 minutes with a very similar spread. Across racial/ethnic groups, only Blacks had a statistically significant change in median LOV over time. From 2006 to 2015, the median LOV for Blacks declined at an annual rate of 2.3 minutes [95% CI: 0.6 to 4.0]). Regionally, only the Northeast experienced a statistically significant change, where median LOV increased by an annual rate of 2.0 minutes (95% CI: 0.1 to 3.9). From 2006 to 2015, the median LOV in this region increased from 151 minutes (IQR: 85 to 246) to 181 minutes (IQR: 104 to 296), which indicates a statistically significant rightward shift in the distribution of LOV in that region. The sickest patients experienced statistically significant increases

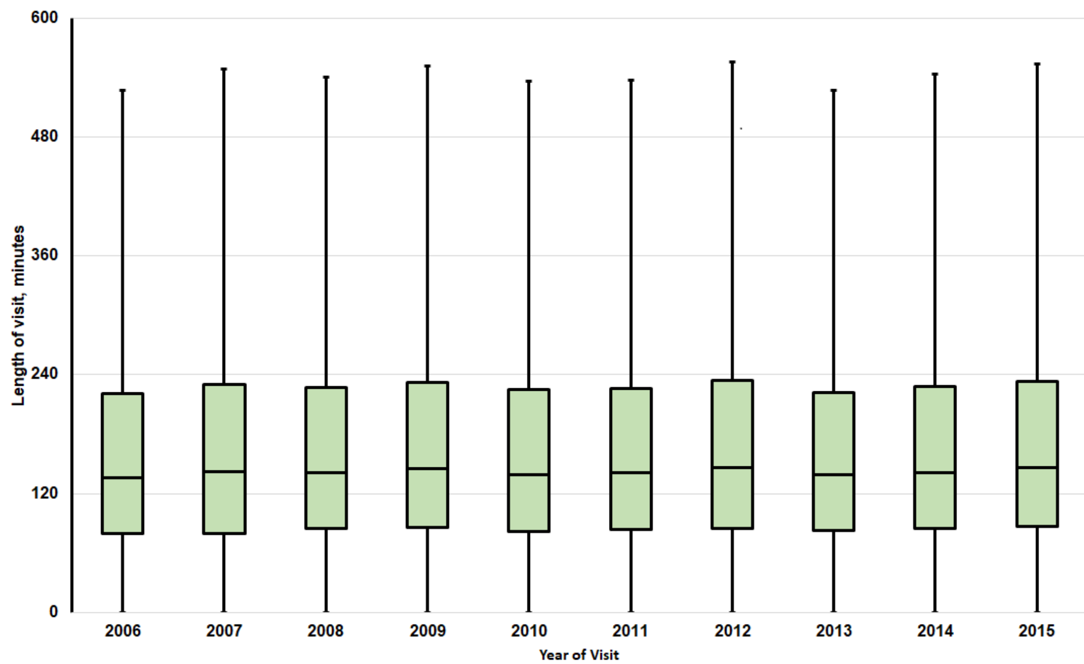

Figure 2.3. Yearly distribution of length of visit (in minutes) for discharged patients, 2006−2015. **Note:** NHAMCS data were weighted to be nationally representative.

in median LOV. It increased at an annual rate of 6.4 minutes (95% CI: 3.9 to 8.9) for the emergent group, and by 2.7 minutes (95% CI: 1.1 to 4.2) for the urgent group. On the contrary, the median LOV decreased for the two least acute groups. It decreased by an annual rate of 3.0 minutes (95% CI: 1.7 to 4.3) for the semi-urgent group, and by 4.8 minutes (95% CI: 2.8 to 6.8) for the non-urgent group. There was no statistically significant change in LOV across the other subgroups. Other variables by which the patient length of visit was characterized are work shift, weekday (yes/no), season, and whether a physician ordered diagnostic services. These results are available in Table A.3 in the Appendix.

2.4.5 Recommended Waiting Time by Acuity

Results about performance on the recommended wait time for the four acuity levels are in Table **??**. Over time, the proportion of visits seen within the respective recommended waiting times increased from 75.5% (95% CI: 72.7% to 78.3%) to

80.8% (95% CI: 77.2% to 84.4%). A similar trend is observed for the two middle acuity groups – the urgent and semi-urgent. The proportion of the urgent group seen within the recommended 60 minute wait time increased from 76.3% (95% CI: 73.1% to 79.5%) to 79.5% (95% CI: 75.1% to 83.9%), and that of the semi-urgent group (to be seen within the recommended 120 minute wait time) increased from 84.7% (95% CI: 81.9% to 87.5%) to 93.7% (95% CI: 92.1% to 95.3%). However, the proportion of the emergent group seen within the recommended wait time decreased from 48.0% (95% CI: 41.6% to 54.4%) to 44.5% (95% CI: 36.5% to 52.5%), though the confidence intervals suggest an insignificant difference in the distribution.

Table 2.5. Percent of Visits in which Wait Time to Consult the ED Practitioner was within the Recommended Time Frame, Stratified by Acuity Level, 2006-2016

	Proportion of visits (with 95% CI) in which wait time was within the recommended time				
Acuity Level	**2006**	**2016**	**Absolute Change**	**OR (95% CI)**	**Adjusted OR (95% CI)**
All visits	75.5 (72.7 to 78.3)	80.8 (77.2 to 84.4)	5.3	1.07 (1.04 to 1.09)	1.07(1.04 to 1.09)
Emergent	48.0 (44.0 to 52.0)	44.5 (36.5 to 52.5)	−3.5	1.00 (0.97 to 1.03)	1.00 (0.97 to 1.03)
Urgent	76.3 (73.1 to 79.5)	79.5 (75.1 to 83.9)	3.2	1.05 (1.05 to 1.08)	1.05 (1.03 to 1.08)
Semi-urgent	84.7 (81.9 to 87.5)	93.7 (92.1 to 95.3)	9.0	1.14 (1.10 to 1.18)	1.13 (1.10 to 1.16)
Non-urgent	100 (0.0)	100 (0.0)	0	1.03 (0.99 to 1.08)	1.03 (0.99 to 1.08)

Notes: Abbreviations: CI, confidence interval. Subgroups may not sum to total visits owing to missing triage values. NHAMCS data were weighted to be nationally representative. Definitions: Emergent: $<$ 15 minutes; urgent: $\leq$ 60 minutes; semiurgent: $\leq$ 2 hours; nonurgent: $\leq$ 24 hours. Unadjusted odds of being treated within the recommended triage time between 2006 and 2016, calculated from an unadjusted survey-weighted logistic regression model. Adjusted odds of being treated within the recommended triage time between 2006 and 2016, calculated from an unadjusted survey-weighted logistic regression model. Adjusted covariates were age, sex, race/ethnicity, source of payment, region, metropolitan status, resident intern involvement, and psychiatric reason for visit.

In adjusted models, there was no statistically significant change in odds of seeing the emergent group in under 15 minutes (AOR: 1.00 [95% CI: 0.97 to 1.03]). However, the odds did increase at lower acuity levels. The odds of seeing the urgent group within 60 minutes of arrival increased over time (1.05 [95% CI: 1.03 to 1.08]). Similarly, that of the semi-urgent group increased over time (1.13 [95% CI: 1.10 to 1.16]). There was no change in the odds for the non-urgent group. Overall, the adjusted odds that any patient with any assigned acuity level was seen within the recommended waiting time increased at an annual rate of 7% (95% CI: 4% to 9%) during the period.

2.4.6 Leaving Without Being Seen (LWBS)

The trend in the proportion of LWBS and the total number of visits from 2006 to 2016 are shown in Figure 2.4. As can be seen from the figure, the number of visits took an upward trend while the proportion of LWBS took a downward trend. The logistic regression model results indicated that the odds of LWBS during this period declined at an annual rate of 6% (95% CI: 4% to 9%).

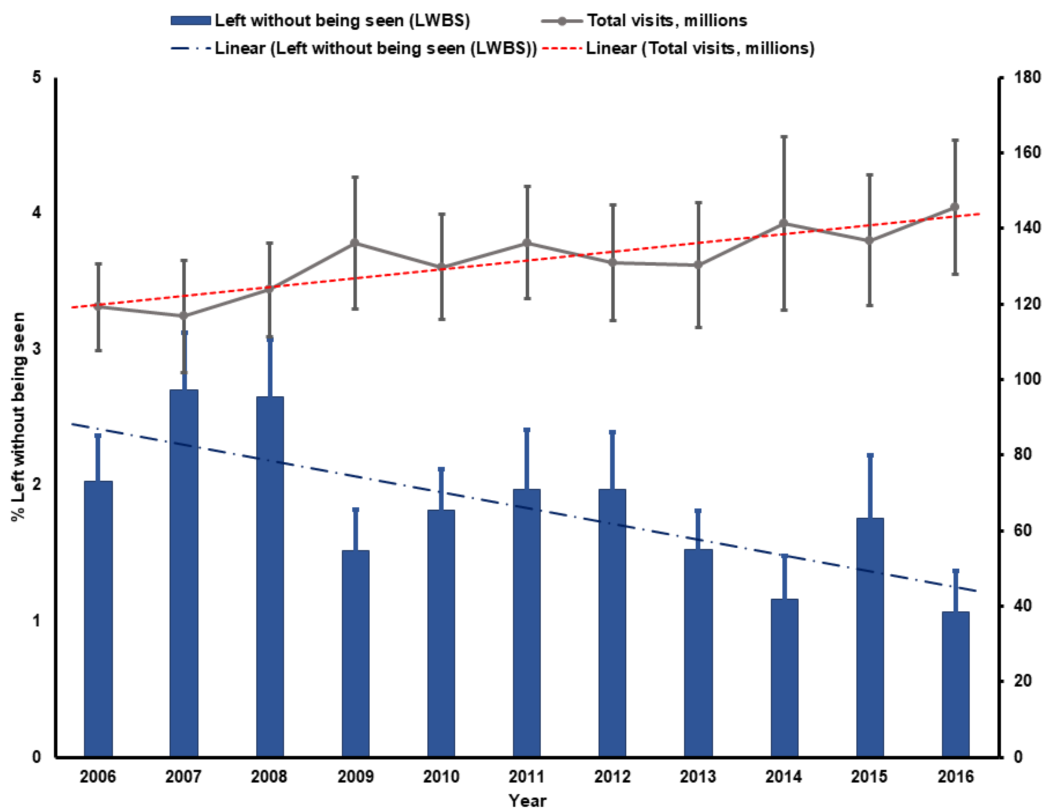

Figure 2.4. Yearly total number of visits and proportion of patients leaving without being seen, 2006−2016.
Note: NHAMCS data were weighted to be nationally representative. Vertical bars are 95% confidence intervals.

2.4.7 Boarding Time

Figure 2.5 indicates that median boarding time was stable though the number of admits declined from 17.1 million (95% CI: 14.3 to 19.9) to 12.3 million (95% CI: 9.9 to 14.7) over time. From 2009 and 2015 (excluding 2012), annual median boarding times were 78 minutes (IQR: 35 to 145), 80 minutes (IQR: 37 to 144), 71 minutes (IQR: 36 to 130), 72 minutes (IQR: 28 to 150), 69 minutes (IQR: 28 to 139), and 78 minutes (IQR: 37 to 148), respectively. Across all years, the largest non-outlier boarding time was 384 minutes (or 6.4 hours).

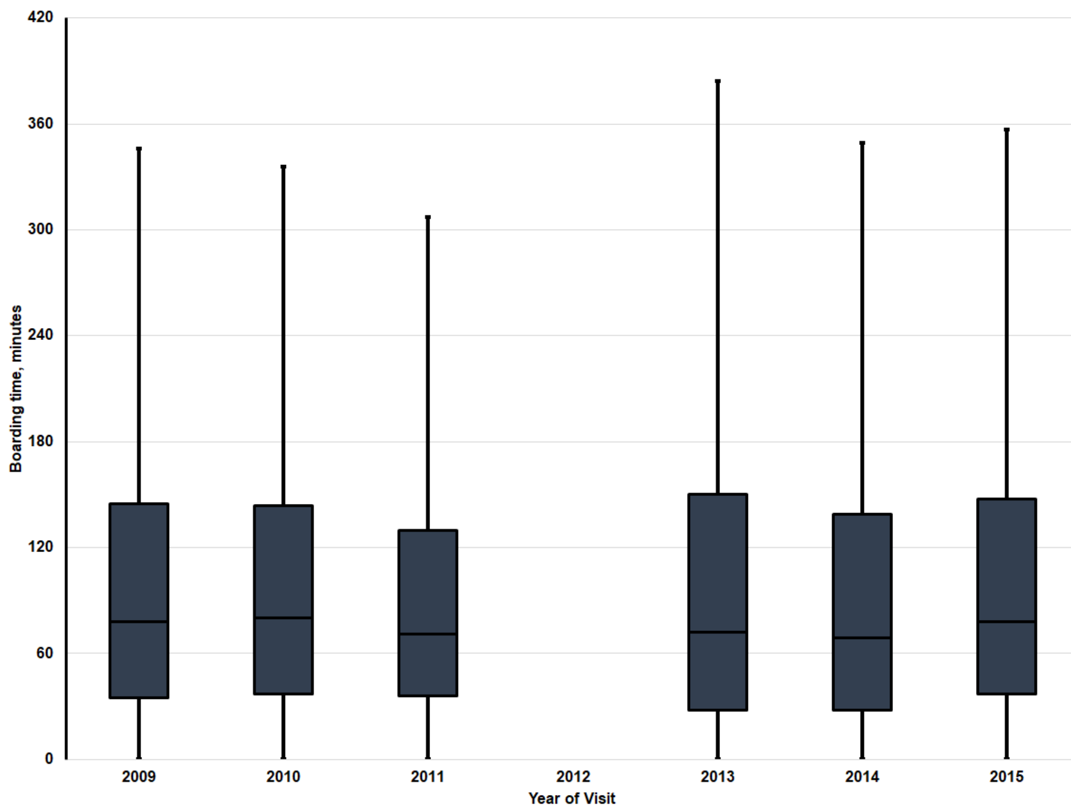

Figure 2.5. Yearly distribution of boarding time (in minutes) for all visits, 2009-2015.
Note: NHAMCS data were weighted to be nationally representative. 2012 data not reported.

2.4.8 Disposition within Four, Six, and Eight Hours

The proportions of visits dispositioned within 4, 6, and 8 hours between 2006 and 2015, stratified by acuity, are shown in Table ??. The odds of being admitted within the three specified hours remained statistically unchanged across all acuity levels. Even though the two lower acuity levels appeared to have had increased odds of admission within 6 hours, the confidence intervals for the odds are very wide. They do not meet the inclusion criteria (standard error greater than 30% of the point estimate) specified by NCHS.

For discharged visits, there were changes over time. The proportion of the emergent, as well as the odds of the emergent being dispositioned with the three specified hours, decreased over time. Within 4, 6, and 8 hours, the proportions for this group declined from 76.2% (95% CI: 72.4% to 80.0%) to 60.3% (95% CI: 53.3% to 67.3%); 90.4% (95% CI: 88.2% to 92.6%) to 80.2% (95% CI: 76.2% to 84.2%); and 95.1% (95% CI: 93.9% to 96.3%) to 89.6% (95% CI: 86.8% to 92.4%), respectively. And the odds of discharge within the three target hours declined by annual rates of 7% (95% CI: 5% to 9%), 7% (95% CI: 4% to 9%), and 7% (95% CI: 4% to 10%), respectively.

However, the two middle acuity levels had increased proportions as well as odds of discharge within 4 hours, but no statistically significant changes within 6 and 8 hours. The proportion of the semi-urgent group discharged within 4 hours increased from 80.7 (95% CI: 77.3% to 83.7%) to 86.4% (95% CI: 83.6% to 89.2%) and that of the non-urgent group increased from 82.7% (95% CI: 78.5% to 86.9%) to 90.1% (95% CI: 83.9% to 96.3%). The odds of discharging the semi-urgent group within 4 hours increased at an annual rate of 5% (95% CI: 2% to 8%). And the odds of discharging the non-urgent group within 4 hours increased at an annual rate of 8% (95% CI: 4% to 12%). The odds for the 6-hour time frame are wider and do not also meet the inclusion criteria set forth by NCHS.

Table 2.6. Percent of Visits in which Wait Time to Consult the ED Practitioner was within the Recommended Time Frame, Stratified by Acuity Level, 2006-2016

			Proportion of visits (SE)		Odds Ratios (95% CI)	
Disposition	**Within**	**Acuity Level**	**2006**	**2015**	**Unadjusted***	**Adjusted****
Admitted	4 h	All	45.5 (2.0)	39.8 (2.5)	0.99 (0.96 to 1.01)	0.99 (0.96 to 1.01)
		Emergent	50.0 (3.3)	40.8 (4.1)	0.97 (0.94 to 1.01)	0.98 (0.94 to 1.01)
		Urgent	43.5 (2.5)	32.7 (3.2)	0.99 (0.96 to 1.02)	0.98 (0.96 to 1.01)
		Semi-urgent	36.8 (3.2)	47.1 (8.1)	1.03 (0.98 to 1.09)	0.98 (0.95 to 1.01)
		Non-urgent	32.1 (6.4)	——‡	——%	——%
	6 h					
		All	71.6 (1.9)	68.0 (2.8)	1.00 (0.96 to 1.01)	1.00 (0.97 to 1.02)
		Emergent	75.2 (2.5)	69.5 (3.8)	0.99 (0.95 to 1.02)	0.99 (0.96 to 1.03)
		Urgent	71.7 (2.3)	61.6 (4.0)	0.99 (0.96 to 1.03)	0.99 (0.95 to 1.03)
		Semi-urgent	63.8 (2.9)	77.7 (6.7)	1.06 (1.00 to 1.12)	1.06 (1.01 to 1.11)
		Non-urgent	51.4 (7.6)	60.9 (10.4)	1.09 (0.99 to 1.20)	1.13 (1.01 to 1.26)
	8 h					
		All	85.4 (1.4)	82.4 (2.4)	0.99 (0.96 to 1.02)	0.99 (0.96 to 1.02)
		Emergent	85.6 (1.9)	82.4 (3.6)	0.99 (0.94 to 1.03)	0.99 (0.95 to 1.03)
		Urgent	86.8 (1.9)	77.3 (3.4)	0.97 (0.93 to 1.01)	0.96 (0.92 to 1.00)
		Semi-urgent	81.3 (2.4)	90.6 (3.9)	1.07 (0.99 to 1.14)	1.06 (0.99 to 1.14)
		Non-urgent	78.1 (6.5)	69.9 (13.8)	0.99 (0.87 to 1.13)	1.01 (0.88 to 1.17)
Discharged	4 h					
		All	78.7 (1.1)	76.4 (1.4)	0.99 (0.97 to 1.02)	1.00 (0.98 to 1.02)
		Emergent	76.2 (1.9)	60.3 (3.5)	0.91 (0.89 to 0.94)	0.93 (0.91 to 0.95)
		Urgent	77.3 (1.6)	66.7 (1.9)	0.96 (0.94 to 0.98)	0.98 (0.96 to 1.00)
		Semi-urgent	80.7 (1.5)	86.4 (1.4)	1.05 (1.03 to 1.08)	1.05 (1.02 to 1.08)
		Non-urgent	82.7 (2.1)	90.1 (3.1)	1.10 (1.04 to 1.17)	1.08 (1.04 to 1.12)
	6 h					
		All	91.5 (0.6)	90.6 (0.8)	0.99 (0.97 to 1.02)	1.00 (0.98 to 1.02)
		Emergent	90.4 (1.1)	80.2 (2.0)	0.91 (0.88 to 0.94)	0.93 (0.91 to 0.96)
		Urgent	91.6 (0.7)	87.1 (1.3)	0.97 (0.95 to 1.00)	0.99 (0.97 to 1.01)
		Semi-urgent	92.1 (0.8)	95.4 (0.6)	1.05 (1.01 to 1.09)	1.04 (1.01 to 1.08)
		Non-urgent	92.6 (1.2)	96.8 (0.8)	1.09 (1.04 to 1.15)	1.07 (1.02 to 1.12)
	8 h					
		All	95.8 (0.3)	95.2 (0.5)	0.98 (0.96 to 1.01)	0.99 (0.97 to 1.01)
		Emergent	95.1 (0.6)	89.6 (1.4)	0.90 (0.87 to 0.93)	0.93 (0.90 to 0.96)
		Urgent	96.2 (0.4)	94.5 (0.6)	0.98 (0.95 to 1.01)	0.99 (0.96 to 1.02)
		Semi-urgent	96.4 (0.4)	97.5 (0.4)	1.03 (0.99 to 1.07)	1.02 (0.98 to 1.06)
		Non-urgent	96.2 (0.6)	97.7 (0.6)	1.06 (0.99 to 1.12)	1.04 (0.98 to 1.10)

Notes: Data were weighted to be nationally representative.
Unadjusted odds of being treated within the specified hour, calculated from survey-weighted logistic regression model. Definition: Admission time was when the order was written; discharge time was when the patient departed the ED.42 Emergency: less than 15 minutes; urgent: less than or equal to 60 minutes; semiurgent: less than or equal to 2 hours; nonurgent: less than or equal to 24 hours.
*Adjusted odds of being treated within the specified hour, calculated from an adjusted survey-weighted logistic regression model. Adjusted covariates were age, sex, race/ethnicity, source of payment, region, metropolitan status, weekday, work shift, season, psychiatric reason for visit, and provision of diagnostic testing/scanning.
‡ Data did not meet inclusion criteria as determined by the National Center for Health statistics.
% Not reported because 2015 estimate for nonurgent group did not meet inclusion criteria.

2.4.9 Study Limitations

This study has some limitations. First, this analysis is reliant on the accuracy of the NHAMCS. The national representativeness of data rests squarely on the representativeness of the sample of EDs. Thus, there is the possibility that the sample of EDs is not perfectly representative of national ED experience. However, the large sample of EDs will likely attenuate any potential bias.

Time measures are particularly prone to misclassification bias and item nonresponse, especially since these data are abstracted from clinical records (electronically since 2012) with many opportunities for error. These biases are likely more severe among the sickest patients, for whom the priority is to provide timely care, and not to document accurate timestamps. However, the accuracy of data and in particular time measures in NHAMCS are taken seriously. Starting in 2005, wait time records were manually checked for consistency, and corrections effected where necessary [39]. Also, quality checks are performed on 10% of the full sample to ensure reliability.

Another limitation concerns the rate of missingness of the boarding time measure, where the response rate was not as high as those of the other outcome variables. To ensure data missingness did not significantly bias the results on boarding time, the distributions of the analytical and missing samples were compared, and it was observed that they did not differ systematically Table A.4 in Appendix A.

Our conservative box-and-whisker plot also indicates that 1 in 20 data points lies outside the "normal" range of most of the data points. But this rate was stable over the study period which suggests that the main results are not affected by potentially extraneous data points. Over time, the proportion of data points outside the "normal" range of most data remained statistically unchanged for all measures, except for the length of visit for admitted patients. However, the large confidence intervals in the latter part of the study period renders this increase insignificant given the large overlap in the 95% confidence intervals in these rates (Figure A.1).

Triage classification scales are subjective and vary among EDs in the US, with a complex rescaling scheme used by NCHS to harmonize values across EDs. Also, approximately 15% of the data did not have a triage level recorded. However, while rescaling triage scores (which is done for less than 10% of records) may affect the quality of data, the subsequent quality checks performed by NCHS may reduce any potential biases. Also, the reliability of triage assessment has been found to be fair to excellent in previous assessments, and a response rate of 85% is high [50, 56].

Finally, several pertinent ED variables such as volume, number of staff, and others could not be adjusted for in the models because they were not reported. As a result, the impact of these variables could not be ascertained.

2.5 Discussion

This study assesses throughput trends U.S. EDs from 2006 to 2016. Overall, there was improvement. During the 11 years, median wait time as well as the proportion that left without being seen nearly halved, and the proportion of patients seen by a qualified practitioner within recommended wait times increased marginally. Over time, increased proportions of low acuity patients were discharged promptly, and previously observed racial and ethnic disparities in wait time disappeared.

These findings contrast favorably with observations in the past. Wilper et al. found that between 1997 and 2004, median wait time increased by approximately 36.4% - which updates to 40.9% as at the end of 2006 [41]. Horwitz and colleagues also found that between 1997 and 2006, the proportion of ED patients seen within recommended wait times declined by 0.8% annually for all patients, and by 2.3% for the sickest patients [40]. The results of this study indicate a marginal increase in this proportion for all patients and no change for the sickest patients. While the latter, in itself, is not an improvement, it is, when compared to the 2.3% decline in the decade prior. Besides, this could be due to challenges related to capturing accurate timestamps for the sickest patients for whom the goal is to provide timely care and not to document accurate data.

There was also improvement in patient boarding time. Between 2009 and 2015, the median boarding time was under 1.5 hours, and about 75% of admitted patients departed the ED within 2.5 hours of being boarded. These rates also compare favorably to findings before 2006 when it was observed that about 20% of EDs reported an average boarding time of 8 hours [57]. In fact, other assessments revealed that before 2006, it was not uncommon for crowded EDs to board patients for more than 48 hours [58].

In the absence of federal or state specified time benchmarks for dispositioning patients, the data reveal that U.S. EDs dispositioned sizable proportions of patients within 8 hours of arrival during the eleven years. As at the end of 2015, approximately 82% of patients were admitted to ward within 8 hours, and approximately 95% were discharged in that time frame. In some developed countries where stringent time benchmarks have been implemented, studies have found improvements in timeliness [59, 60]. However, other studies have documented high levels of workload-related stress and decreased morale among ED staff following the implementation of such rules [61], with possible consequences for patients, as predicted by experts [62, 63]. The 98% target disposition within 4 hours implemented in the UK was subsequently revised to 95%, and that of Western Australia from 90% to 85% due to these concerns [64].

Worthy of note is that while there were improvements in wait time across all acuity levels, there was an increase in the length of visit for the sickest patients that were discharged the same day. The data reveal that, over time, decreased proportions of the sickest patients were discharged within 4, 6, and 8 hours of arrival. This could be due to residual differences in patient characteristics or increased ancillary testing or specialty consultation before discharging the acutely ill, though patient and ED characteristics including age, sex, race/ethnicity, source of payment, region, metropolitan status, weekday/weekend, time of arrival, season, and the provision of diagnostic services during the visit were adjusted for in the models. Also worthy of mention is that CMS had announced that starting in October 2020, median wait time would be excluded from the reported ED timeliness measures, citing concerns raised by a technical team of experts [65]. While this study did not assess whether or not the observed improvements in wait time reflect the impact of the CMS initiative, it remains a strong possibility. What the removal of this measure holds for the future remains uncertain. These results are expected to prompt a careful re-assessment of that decision.

A potentially contradictory result is that the observed improvements were attained in the face of a rising trend in the volume of ED visits. Between 2006 and 2016, ED presentations increased by 22.1%, nearly thrice the U.S. population growth during that period [55]. The logical expectation is that increased visits will impede throughput. However, some have argued that slow throughput is largely the result of how patients are processed within the ED, rather than the influence of "input" factors such as patient volume [2, 66, 67, 68]. Several studies have demonstrated that making changes to aspects of ED processing such as registration, data processing, triage, diagnostics, and work assignment procedures could significantly enhance speedy throughput [6, 69, 70, 71, 72, 73, 74].

There are a few possible interrelated reasons for the observed improvement in timeliness despite the increased volume of visits. First, the recent digital revolution in the health care space may partly explain this result. It has been argued that health information technology (HIT) can eliminate redundancies in care processes, increase productivity, and facilitate management and coordination of patient flow [1]. The implementation of the Health Information Technology for Economic and Clinical Health Act (HITECH) in 2009 created incentives for the widespread adoption and use of HIT by health care providers. The use of these systems in the health care space is now commonplace. Selck and Decker found that between 2007 and 2010, the use of advanced EHR technology in U.S. EDs increased almost tenfold [3]. Further, they found that patients that visited EDs that used advanced EHR systems waited, on average, 6 minutes less to see the physician when compared to those that visited EDs with only basic EHR systems.

Earning financial rewards and avoiding sanctions by payers may have influenced the implementation of some throughput reforms. For instance, the recent inclusion of ED timeliness measures to CMS pay-for-reporting initiative may have incentivized providers to pay attention to these measures [75]. Hospitals that report these measures earn a reward with increased payments and those that fail to report earn reduced payments. EDs are also increasingly implementing initiatives to achieve faster throughput and increase patient satisfaction. Some EDs also advertise average wait times on various media to attract patients [76]. Throughput strategies such as bedside registration, self-check-in kiosks, "immediate bedding," "fast track," "physician-in-triage", and other similar throughput initiatives have been implemented in many EDS today [77, 78, 79]. These throughput-enhancing mechanisms could potentially improve patient flow despite increased visits.

2.6 Conclusion

Overall, the results indicate there was improvement in ED throughput over time. In particular, there were decreases in patient wait time, patients leaving without meeting the emergency physician, and boarding time, especially when compared to the previous decade. However, the time to discharge did increase for the sickest patients. This could be due to increased use of diagnostic services to reduce uncertainty before discharging acutely ill patients. Nevertheless, stakeholders must take steps to ensure that throughput reforms intended to improve ED timeliness do not negatively impact any subgroup of patients.

CHAPTER 3
RACIAL AND ETHNIC DISPARITIES IN PATIENT WAIT TIME

"Of all the forms of inequality, injustice in health is the most shocking and the most inhuman because it often results in physical death" – Dr. Martin Luther King Jr.

3.1 Introduction

As previously described, equity and timeliness are two of six attributes of a quality health care system as defined by the National Academy of Medicine (NAM) [1]. Taken together, the two attributes dictate that personal characteristics should not have a bearing on the receipt of timely and quality care. Yet, previous evidence suggested that Black and Hispanic patients waited significantly longer than their White counterparts to receive treatment in U.S. emergency departments [80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91].

Recent broad-based efforts targeted at improving health care quality may reduce or eliminate disparities. For instance, using publicly-reported data, one study assessed racial and ethnic disparities in quality of care for patients hospitalized for acute myocardial infarction, heart failure, and pneumonia and found improvements in quality for all racial and ethnic groups as well as reductions in disparities [92]. Another study assessed racial disparities using publicly reported quality-of-care measures for Medicare beneficiaries and found substantial improvements in quality for all and a narrowing of disparities in most of the measures [93]. Other studies have documented improvements in equity following implementation of similar broad-based quality initiatives [94]. While in general, such initiatives may reduce, worsen, or have no effect on equity, these recent findings support the proposition that broad-based improvements in quality of care are associated with improvements in equity [95].

Several recent initiatives may improve timeliness in U.S. EDs and reduce or eliminate previously observed disparities in patient wait time. For instance, the

American College of Emergency Physicians (ACEP) has proposed the adoption of "high-impact" solutions for reducing crowding in EDs [96]. Starting in October 2013, the Centers for Medicare and Medicaid Services (CMS) collected and publicly-reported median ED wait time data [97]. In general, the passage of the Patient Protection and Affordable Care Act (PPACA) in 2010 ushered in a new era of health care delivery in the US, emphasizing accountability and quality care delivery [98].

The goal of this study is to assess trends and sources of ED wait time disparities across racial and ethnic groups by analyzing more recent national data. The study also analyzes changes in the use of proposed throughput strategies targeted towards improving throughput in U.S. EDs.

3.2 Literature Review

There is enormous evidence that point to wide racial and ethnic disparities in U.S. emergency department wait time. The evidence suggests that Blacks and Hispanics, children and adults alike, wait significantly longer than their white counterparts to access care in U.S. emergency departments.

Bradley et al. estimated that the average door-to-balloon time (i.e., time from arrival to first treatment of clogged arteries using balloon therapy) difference was 19 minutes between Blacks and Whites, and 11 minutes between Hispanics and Whites. These differences were significant even after controlling for time, patient, and provider heterogeneity [80]. The study further showed that the time from arrival to getting drugs was significantly longer for Hispanics, and more so for Blacks. Lopez et al. also previously found that black and Hispanic patients that presented to the ED with chest pain were less likely to be offered treatment promptly compared to white patients [81].

Sonnenfeld and colleagues [82] analyzed wait times by the proportion of minorities at a given ED and found that there was a 23% (unadjusted) and 13%

(adjusted) increase in wait time for every 25 point increase in the percent of black patients treated at an ED. Though they found some evidence of within-ED difference in wait time between Whites and Blacks, they concluded that the differences were as a result of Blacks accessing high-volume EDs which turn to have long wait times. Like Sonnenfeld and colleagues, Pines et al. [83] analyzed nationally representative data and found that, the Black race independently predicted wait time after adjusting for gender, metropolitan status of ED, and disease severity, they observed that Black patients had a 19 percent greater odds of waiting more than 30 minutes when compared to Whites.

Wilper et al. [84] evaluated changes in ED wait time from 1997 to 2004 for patients aged eighteen or older. They found that during the said period, median wait time increased by an annual average of 4.1%. They narrowed the analysis down to patients diagnosed with acute myocardial infarction and found that median wait time increased by 11.2% per annum. They further found that Black race and Hispanic ethnicity were significant predictors of prolonged waiting.

Ray et al. [85] analyzed time patients spent traveling to, waiting for, and receiving ambulatory medical care using the American Time Use Survey (ATUS) data from 2005 to 2013 for respondents 18 years and older. They concluded that time spent seeking care was significantly longer for racial and ethnic minorities.

Haywood et al. [86] undertook a cross-sectional analysis of the National Hospital Ambulatory Medical Survey data for years 2003 to 2008. They focused on patients presenting with the sickle cell disease (SCD). They found that average wait time was 25% longer for patients with SCD than for the general patient pool. Further, they observed that Black race was a significant predictor of prolonged waiting.

Park et al. [87], James et al. [88], and Zhang et al. [89] all focused on wait time for children (under 15, 16 and 18 years, respectively). All studies analyzed the same data, the National Hospital Ambulatory Medical Care Survey. James et al. analyzed data for the period 1997 to 2000, and concluded there were significant disparities in ED wait time between White and minority children. They alluded to "discrimination, cultural incompetence, language barriers, and other social factors" as the reasons for the disparities. Park analyzed data for years 2005 and 2006 and concluded that there were "sizeable" differences in ED wait time between White children and children of minorities, and that this difference was evident both within and across EDs. Zhang analyzed more recent data, from 2005 to 2016, and concluded that there were significant differences in waits between Whites and minorities (Blacks and Hispanics) but did not investigate the source of the disparity.

Qiao et al. [90] also analyzed the National Hospital Ambulatory Medical Care Survey for year 2008. Having analyzed 34,143 visiting at 353 EDs, they found that, on average, Black patients waited 16 minutes longer than their White counterparts. They noted that these differences were only significant among low acuity patients but not the sickest patients.

Karve et al. [91] analyzed ED wait time for adult patients presenting with stroke using data from the National Hospital Ambulatory Medical Care Survey for years 1997–2000 and 2003–2005. In adjusted models, they concluded that there were significant differences in ED wait time between Whites and Blacks but that there was no difference between Whites and Hispanics.

3.3 Methods

3.3.1 Study Design and Data Source

As in the previous chapter, this was a cross-sectional analysis of nationally representative data from the National Hospital Ambulatory Medical Care Survey (NHAMCS). Data from 2003 to 2017 was used for this analysis. Details about NHAMCS data collection and processing procedures are described in Section 2.3.2. Further and more detailed description of NHAMCS data is also available in the data documentation provided by the National Center for Health Statistics (NCHS) [49].

3.3.2 Outcome Measure

The primary outcome variable was ED wait time (in minutes), defined in NHAMCS as the time between arrival and first contact with a qualified practitioner (physician, physician assistant, or nurse practitioner). The sample includes patients of all ages who presented to the ED from 2003 to 2017. Of the 433,684 sampled encounters, wait time was recorded for 359,992 (83% response rate), a high response rate for a survey of this nature [50]. The sample distributions of the analytical and missing samples were compared, and it was found that they did not differ systematically. These results are available in Table B.1 in Appendix B. Missing data were not imputed.

3.3.3 Independent Variables

The primary independent variable was race/ethnicity grouped into non-Hispanic White, non-Hispanic Black, and Hispanic. Starting in 2009, the race/ethnicity variable in NHAMCS was coded as a 4-level variable that identified (1) White, Non-Hispanic; (2) Black, Non-Hispanic; (3) Hispanic, (4) Non-Hispanic Other. Even with the merger of all other racial categories (i.e. Asian Only; Native Hawaiian, Other Pacific Islander Only; American Indian/Alaska Native; Multiple Races) into "Non-Hispanic Other", the annual sample sizes for this group was still very small. As a result, subgroup analyses generally yielded estimates with large standard errors (thus, likely non-representative of population). This group was excluded from the analysis.

Other variables were age (0-17, 18-44, 45-64, 65+ years), sex, source of payment (private, Medicare, Medicaid/CHIP, self-pay, Other), mode of arrival (ambulance, other transport), acuity level (emergent [a merger of acuity levels 1 and 2], urgent [level 3], semi-urgent [level 4], non-urgent [level 5]), and primary reason for visit (Schneider et al. classification [99]. ED variables were region (Northeast, Midwest, South, West), metropolitan status (metropolitan, non-metropolitan), and resident/intern physician involvement (yes, no). Time-based variables were work shift (07:00−15:00, 15:00−00:00, 00:00−07:00), weekday (yes, no), season (winter, spring, summer, autumn), and year of visit.

3.3.4 Statistical Analysis

Trends in median wait time were assessed using joinpoint statistical analysis (National Cancer Institute Joinpoint statistical software Version 4.8.0.1 [100]). The joinpoint method identifies statistically significant transition points on a trend line and fits piecewise-linear regression models to the segments separated by the transition points (called joinpoints). The recommendations of the National Center for Health Statistics followed in this analysis [101], using the Joinpoint software to identify the number and location of joinpoints, and then the Survey Package [53] to estimate annual changes in the median (by fitting survey-weighted piecewise regression models through the median). This analysis was subset by racial/ethnic group. Trends in annual median wait time difference between Whites and the two minority groups were also assessed stratified by patient acuity level [90].

Unadjusted survey-weighted bivariate linear regression models were used to assess trends in wait time across patient subgroups. Importantly, trends were assessed for patients presenting with select time-critical chief complaints that are readily recognizable at triage and should trigger prompt response from practitioners – patients presenting with chest or heart pain; lung congestion, chest cold, lung pain, or breathing problems; stomach or abdominal pain, cramps, and spasms; and slurring, diminished vision, numbness, confusion, neurologic weakness, nerve block; and injury.

To assess the source of disparities, survey-weighted log-linear regression models were fit to examine the relationship between wait time and race/ethnicity while controlling for all the aforementioned independent variables and covariates. ED fixed effects were added to this adjusted "baseline" model to ascertain the source of disparities. With the inclusion of ED identifiers, any observed racial/ethnic disparity in the baseline model should not change significantly if the disparity is within EDs. Otherwise, the disparity is across EDs. Because ED identifiers in NHAMCS are not unique to sampled EDs, models were fit to data for each year to compare the baseline model to the fixed-effects model.

In secondary analysis, trends in the percent of patients seen at EDs that indicated they used certain strategies/technology were assessed. Starting in 2007, NHAMCS fielded questions about EDs use of: (1) separate fast-track units for non-urgent care; (2) bedside registration; (3) "pool" nurses (i.e., nurses called in to respond to surges in demand) (4) electronic dashboards (display updated patient information and integrates multiple data sources); (5) fully implemented electronic health record (EHR) systems; and starting in 2012, (6) physician/practitioner at triage; (7) immediate bedding (no triage when ED is not at capacity). Joinpoint software was used to identify the number and location of joinpoints, and a survey-weighted linear probability model used to assess trends in the percent of patients stratified by race/ethnicity.

3.4 Results

3.4.1 Sample Characteristics

Table 3.1 shows the distribution of the sample of visits by race/ethnicity. Of the 359,992 visits, 61.4% were non-Hispanic White, 23.3% were non-Hispanic Black, and 15.3% were Hispanic. The sex distribution of patients across groups was similar. Blacks and Hispanics mostly presented with Medicaid/CHIP coverage, while Whites mostly presented with private insurance. The proportion of Whites, Blacks, and Hispanics presenting to EDs located in metropolitan areas was 74.8%, 88.4%, and 87.3%, respectively. The majority of Whites and Blacks were in the South, while Hispanics were mostly in the West. The distribution of patients across acuity levels was very similar across the groups.

Table 3.1. Characteristics of Unweighted Sample of Visits to U.S. Emergency Departments, 2003-2017

Characteristic	**All**	**White**	**Black**	**Hispanic**
N (%)	359,992 (100)	221,138 (61.4)	83,723 (23.3)	55,131 (15.3)
Age group (years), n (%)				
0−17	81,342 (22.6)	41,583 (18.8)	21,060 (25.2)	18,699 (33.9)
18−44	145,401 (40.4)	85,902 (38.8)	36,628 (43.7)	22,871 (41.5)
45−64	77,373 (21.5)	49,794 (22.5)	18,571 (22.2)	9,008 (16.3)
65+	55,876 (15.5)	43,859 (19.8)	7,464 (8.9)	4,553 (8.3)
Sex, n (%)				
Female	195,743 (54.4)	119,607 (54.1)	46,783 (55.9)	29,353 (53.2)
Male	164,249 (45.6)	101,531 (45.9)	36,940 (44.1)	25,778 (46.8)
Payment Source, n (%)				
Private insurance	111,861 (31.1)	80,547 (36.4)	19,043 (22.7)	12,271 (22.3)
Medicare	62,322 (17.3)	47,638 (21.5)	10,160 (12.1)	4,524 (8.2)
Medicaid/CHIP	100,122 (27.8)	46,015 (20.8)	31,677 (37.8)	22,430 (40.7)
Self-Pay	47,814 (13.3)	26,008 (11.8)	13,046 (15.6)	8,760 (15.9)
Other	37,873 (10.5)	11,897 (5.4)	5,036 (6.0)	3,955 (7.2)
Region, n (%)				
Northeast	79,154 (22.0)	47,848 (21.6)	16,091 (19.2)	15,215 (27.6)
Midwest	83,130 (23.1)	58,074 (26.3)	18,791 (22.4)	6,265 (11.4)
South	132,919 (36.9)	74,678 (33.8)	42,620 (50.9)	15,621 (28.3)
West	64,789 (18.0)	40,538 (18.3)	6,221 (7.4)	18,030 (32.7)
Metropolis*, n ()				
Yes	287,536 (79.9)	165,381 (74.8)	74,007 (88.4)	48,148 (87.3)
No	48,631 (13.5)	41,440 (18.7)	4,707 (5.6)	2,484 (4.5)
Acuity level, n (%)				
Emergent	61,219 (17.0)	39,389 (17.8)	13,171 (15.7)	8,659 (15.7)
Urgent	128,283 (35.6)	79,326 (35.9)	29,693 (35.5)	19,264 (34.9)
Semi-urgent	86,224 (24.0)	51,107 (23.1)	21,509 (25.7)	13,608 (24.7)
Non-urgent	30,530 (8.5)	18,116 (8.2)	7,301 (8.7)	5,113 (9.3)
No triage/Unknown	53,736 (14.9)	33,200 (15.0)	12,049 (14.4)	8,487 (15.4)
Primary RFV, n (%)				
Chest/heart pain	18,507 (5.1)	11,936 (5.4)	4,313 (5.2)	2,258 (4.1)
Lung congestion, chest cold, lung pain, breathing problems	4,785 (1.3)	2,969 (1.3)	1,121 (1.3)	695 (1.3)
Stomach or abdominal pain, cramps and spasms	24,265 (6.7)	14,736 (6.7)	5,133 (6.1)	4,396 (8.0)
Injury	124,962 (34.7)	83,001 (37.5)	24,813 (29.6)	17,148 (31.1)
Slurring, diminished vision, numbness, confusion, neurologic weakness, nerve block	3,828 (1.1)	2,514 (1.1)	731 (0.9)	439 (0.8)

Notes: Subgroups may not sum to total N due to missing/unknown responses or sum to 100% due to rounding. *Metropolitan status was not reported in 2012 in the NHAMCS public use file.

3.4.2 Unadjusted Trends in Median Wait Time by Race/Ethnicity

From 2003 to 2017, the median wait time for Whites decreased from 25 to 16 minutes; that of Blacks decreased 32 to 18 minutes; and that of Hispanics decreased from 37 to 17 minutes (P-value for trend < 0.001 for all) (Table 3.2).

Figure 3.1 shows trends in the annual unadjusted median wait time by race/ethnic group. For Whites, median wait time increased by 1.3 minutes annually (95% confidence interval [CI]: 0.7 to 2.0) from 2003 through 2008. It then decreased by 3.0 minutes (95% CI: -4.0 to -2.0) annually from 2008 through 2012. From 2012 to 2017, it declined by 1.7 minutes (95% CI: -2.1 to -1.2). For Blacks, median wait time initially increased by 2.0 minutes annually (95% CI: 0.8 to 3.2) from 2003 through 2008. Then, it decreased by 3.8 minutes annually (95% CI: -4.9 to -2.7) from 2008 through 2015, and remained fairly flat from 2015 through 2017. For Hispanics, the trend in median wait time remained statistically unchanged from 2003 through 2009 (Slope = 0.50 [95% CI: -0.5 to 1.5]). From 2009 to 2012, it decreased by 4.7 minutes annually (95% CI: -7.0 to -2.4). From 2012 through 2017, the decline continued but at a slower rate of 1.5 minutes annually (95% CI: -2.8 to -0.2).

3.4.3 Unadjusted Subgroup Trends

There were decreases in the median wait time across the age and sex subgroups for all three groups ($p-$value for trend < 0.001) (Table 3.2). The trend was similar with respect to source of payment. However, while there were significant declines in median wait time for EDs located in metropolitan areas across all groups, there was no statistically significant change for EDs located in non-metropolitan areas, except for Blacks, who experienced a significant decline (p-values: 0.10, 0.02, and 0.13 for Whites, Blacks, and Hispanics, respectively). This appears to be so because Blacks had longer median wait times in 2003 (15 versus 20 versus 15 minutes for Whites, Blacks, and Hispanics, respectively) and similar in 2017 (15, 17, and 17 minutes for

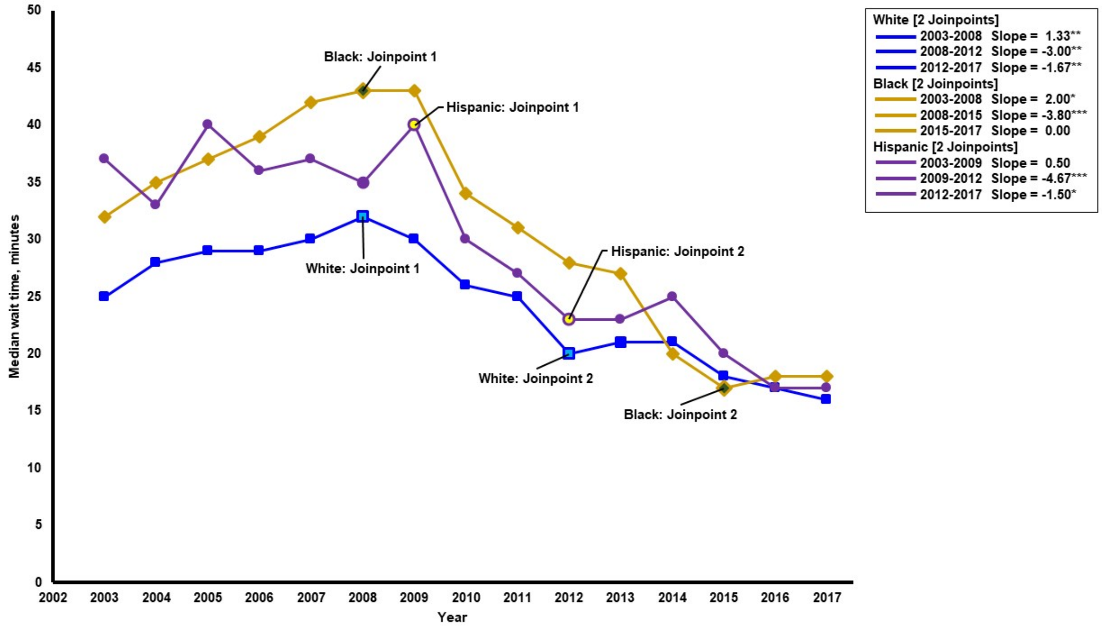

Figure 3.1. Unadjusted annual median ED wait time stratified by race/ethnicity, 2003−2017. Data were weighted to be nationally representative.
Note: Final selected model: White – 2 Joinpoints (2008, 2012), Black – 2 Joinpoints (2008, 2015), Hispanic – 2 Joinpoints (2009, 2012).

Whites, Blacks, and Hispanics, respectively). There were significant changes in the median wait at "majority-White" EDs ($\leq$ 90% White) for all three groups, where median wait time decreased from 21 to 16 minutes for Whites, 35 to 14 minutes for Blacks, and from 33 to 11 minutes for Hispanics.

Table 3.2. Median (IQR) Wait Time and Trends Across Patient Subgroups, Stratified by Race/Ethnicity, 2003-2017

	White			Black			Hispanic		
Characteristic	2003	2017	P-value	2003	2017	P-value	2003	2017	P-value
All	25 (11, 51)	16 (7, 35)	<0.001	32 (14, 70)	18 (7, 44)	<0.001	37 (16, 80)	17 (7, 42)	<0.001
Age group, yrs									
0−17	27 (13, 55)	19 (8, 42)	<0.001	35 (16, 66)	20 (8, 43)	<0.001	40 (18, 80)	22 (10, 53)	<0.001
18−44	26 (12, 53)	15 (7, 35)	<0.001	32 (14, 74)	16 (7, 44)	<0.001	37 (17, 82)	16 (6, 38)	<0.001
45−64	25 (11, 50)	15 (6, 32)	<0.001	33 (15, 77)	17 (6, 48)	<0.001	34 (14, 74)	12 (5, 31)	<0.001
65+	21 (9, 43)	13 (6, 34)	<0.001	25 (10, 57)	16 (5, 39)	<0.001	35 (14, 70)	10 (3, 26)	<0.001
Sex, years									
Female	26 (12, 52)	16 (7, 38)	<0.001	34 (15, 72)	17 (7, 43)	<0.001	36 (17, 82)	18 (7, 41)	<0.001
Male	24 (11, 50)	15 (7, 33)	<0.001	30 (14, 66)	19 (7, 46)	<0.001	37 (15, 75)	16 (7, 43)	<0.001
Payment									
Private insurance	26 (12, 51)	16 (7, 36)	<0.001	32 (14, 68)	23 (7, 51)	<0.001	38 (15, 77)	14 (5, 35)	<0.001
Medicare	22 (10, 45)	14 (6, 35)	<0.001	26 (11, 60)	16 (7, 39)	<0.001	32 (11, 67)	10 (3, 26)	<0.001
Medicaid/CHIP	25 (12, 53)	16 (7, 35)	<0.001	32 (15, 66)	16 (6, 39)	<0.001	32 (11, 67)	10 (3, 26)	<0.001
Self-Pay	26 (12, 54)	16 (8, 41)	<0.001	34 (15, 70)	25 (11, 71)	0.001	38 (18, 82)	20 (9, 51)	<0.001
Other	25 (11, 58)	12 (7, 28)	<0.001	39 (15, 89)	10 (6, 36)	<0.001	38 (17, 77)	13 (5, 30)	<0.001
Metropolis*									
Yes	29 (13, 58)	16 (6, 38)	<0.001	35 (15, 75)	18 (7, 45)	<0.001	39 (17, 82)	17 (7, 43)	<0.001
No	15 (8, 30)	16 (7, 30)	0.10	20 (8, 40)	17 (10, 34)	0.02	15 (10, 40)	17 (7, 28)	0.13
% Black/Hispanic									
<10%	21 (10, 45)	16 (7, 32)	0.03	35 (18, 65)	14 (11, 38)	0.002	33 (15, 74)	11 (6, 28)	0.04
50%+	30 (15, 64)	15 (6, 35)	<0.001	33 (14, 72)	20 (7, 49)	<0.001	44 (19, 90)	16 (7, 38)	<0.001
Acuity Level									
Emergent	25 (13, 45)	15 (6, 36)	0.002	29 (15, 55)	19 (7, 48)	0.30	36 (18, 71)	20 (8, 47)	0.21
Urgent	40 (20, 72)	16 (7, 36)	<0.001	50 (23, 93)	19 (7, 42)	<0.001	51 (24, 90)	20 (8, 47)	<0.001
Semi-urgent	38 (18, 75)	19 (11, 37)	<0.001	53 (25, 105)	16 (8, 42)	<0.001	54 (23, 107)	21 (6, 56)	<0.001
Non-urgent	21 (11, 53)	15 (6, 34)	<0.001	35 (15, 80)	16 (7, 39)	<0.001	33 (15, 79)	11 (4, 28)	<0.001
Primary Reason for Visit									
Chest/heart pain	18 (8, 38)	15 (7, 34)	0.06	24 (12, 58)	16 (7, 52)	<0.001	28 (10, 60)	13 (7, 46)	0.03
Lung congestion, chest cold, lung pain, breathing problems	18 (7, 38)	14 (8, 40)	0.25	23 (13, 48)	10 (4, 18)	0.004	25 (15, 87)	14 (4, 75)	0.08
Stomach or abdominal pain, cramps and spasms	28 (13, 58)	16 (6, 35)	<0.001	40 (16, 83)	20 (7, 52)	<0.001	45 (20, 110)	23 (8, 50)	0.009
Injury	24 (11, 50)	15 (6, 34)	<0.001	30 (13, 65)	18 (6, 47)	<0.001	35 (15, 75)	15 (6, 31)	<0.001
Slurring, diminished vision, numbness, confusion, neurologic weakness, nerve block	26 (10, 59)	19 (9, 41)	0.01	26 (20, 55)	13 (7, 37)	0.002	44 (11, 71)	−−	<0.001

Across acuity levels, Whites experienced significant declines in median wait time ($p-$value for trend < 0.001). Blacks and Hispanics on the other hand experienced significant declines in median wait time for patients triaged as low acuity (p-value for trend < 0.001) but there was no statistically significant change in trend for those triaged emergent (p-value for trend: 0.30 and 0.21, respectively). This might be due to increased attention paid to reducing wait times for lower acuity minorities, which took away time for the sickest patients. It may also be due to inaccuracies in time recorded for the sickest patients who may have received timely treatment without ED staff paying attention to timestamps.

The median wait time for Whites presenting with chest or heart pain decreased from 18 to 15 minutes, with no statistically significant change in trend ($p-$value $=$ 0.06). This appears to be so because White patients with such complaints already experienced shorter wait times in earlier years. Similarly, the trend did not change for Hispanics presenting with such complaints ($p-$value for trend $= 0.03$), but over time, the median wait time for this group decreased from 28 to 13 minutes, lower than that for Whites. On the other hand, the median wait time for Black patients presenting with chest or heart pain decreased from 24 to 16 minutes ($p-$value for trend < 0.001). These results suggest that over time, there were improvements in timeliness and equity in timeliness for patients presenting to the ED with chest or heart pain.

The median wait time for Whites presenting with complaints of lung congestion, chest cold, lung pain, or breathing problems decreased from 18 to 14 minutes, but there was no change in trend ($p-$value $= 0.25$). There was a statistically significant decrease in the median and trend for Blacks ($p-$value $= 0.004$) but not Hispanics ($p-$value $= 0.08$), though both had median values of under 15 minutes in 2017. There were significant declines in the medians for those presenting with injury, stomach or abdominal pain, cramps, and spasms as well as general trends in wait time for all groups.

Figure 3.2 shows trends in annual unadjusted median wait time differences between Whites and the two minority groups stratified by patient acuity level. For Blacks versus Whites, the difference for the emergent and urgent) groups from 2003 to 2017 was near zero, and the trend remained statistically unchanged over time. However, the semi-urgent and non-urgent groups' differences were significantly different from zero between 2003 and 2007. Median wait time for the semi-urgent group during this period increased annually by 3.0 minutes (95% CI: 0.6 to 5.4), then decreased by 3.6 minutes (95% CI: −4.8 to −2.3) between 2007 and 2014. The change after 2014 was not statistically significant (1.9 minutes [95% CI: −1.9 to 5.7]). For the non-urgent group, there was no statistically significant change in the trend from 2003 to 2008 (2.6 minutes [95% CI: −1.3 to 6.5]). However, from 2008 to 2017, there was an annual decline in wait time of 3.0 minutes (95% CI: −4.6 to −1.4).

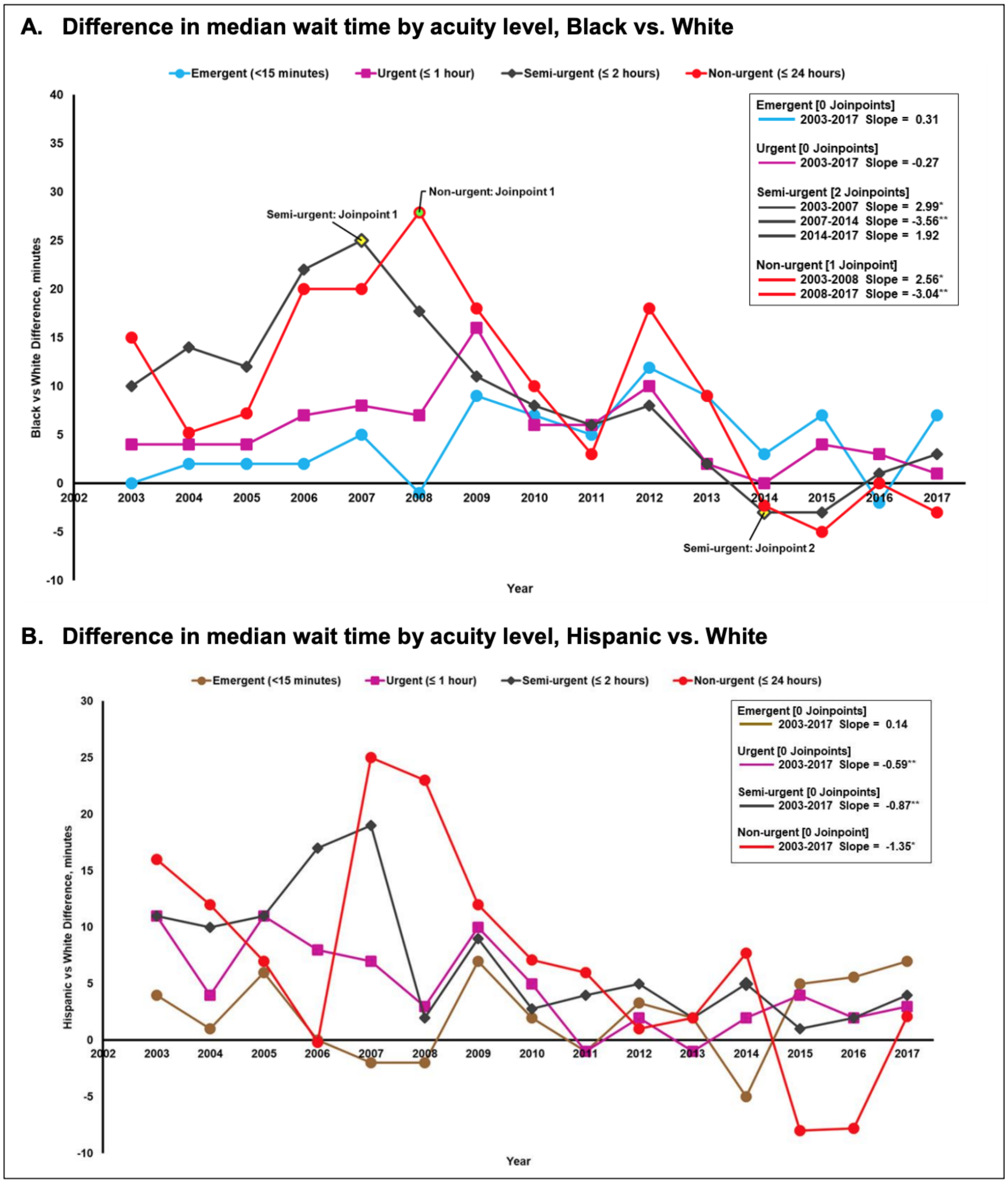

Figure 3.2. Unadjusted differences in median wait time between Whites and the two minority groups (Blacks and Hispanics) stratified by acuity level, 2003−2017.
Notes: Data were weighted to be nationally representative. Final selected model (**Panel A**): Emergent – 0 Joinpoints, Urgent – 0 Joinpoints, Semi-urgent – 2 Joinpoints, Non-urgent – 1 Joinpoint. Final selected model (**Panel B**): Emergent – 0 Joinpoint, Urgent – 0 Joinpoints, Semi-urgent – 0 Joinpoints, Non-urgent – 0 Joinpoints. Statistical significance: ***p<0.001; **p<0.01; *p<0.05.

3.4.4 Adjusted Differences and Sources of Disparities

Except for the year 2005, the annual adjusted baseline models indicate statistically significant differences in wait time between Blacks and Whites from 2003 to 2012 (Table 3.3). On controlling for ED fixed effects, the differences disappear. No statistically significant differences were observed in the baseline and fixed-effects models after 2012.

For Hispanics, the baseline models suggest there were statistically significant differences in wait time (compared to Whites) in the years 2003 and 2005 to 2007. On controlling for ED fixed effects, all differences disappeared except in the year 2003. Nevertheless, the difference in 2003 did diminish by 58.3% on the inclusion of ED fixed effects. No statistically significant differences were observed in the baseline and fixed-effects models after 2007.

From 2003 to 2017, between 5 to 19% of the variability in wait time was explained by the variables in the baseline model, while the fixed-effects model explained 29 to 49% of the variability.

3.4.5 ED Use of Throughput Strategies

Figure 3.3 shows trends in the percent of patients seen at EDs that indicated they used specific throughput strategies from years 2007 to 2017, stratified by racial/ethnic group. The proportion of patients seen at EDs that indicated they used electronic dashboards increased at annual rates of 3.8% (95% CI: 2.9% to 4.7%) for Whites, 3.2% (95% CI: 2.2% to 4.2%) for Blacks, and 3.3% (95% CI: 2.1% to 4.4%) for Hispanics. Similarly, the proportion of patients seen at EDs that indicated they fully implemented EHR systems increased annually by 7.1% (95% CI: 6.3% to 7.8%) for Whites, 5.7% (95% CI: 4.5% to 6.8%) for Blacks, and by 6.8% (95% CI: 5.9% to 7.7%) for Hispanics. The proportion of patients seen at EDs that indicated they practiced "pool" nursing increased annually by 2.1% (95% CI: 1.2% to 2.9%) for

Table 3.3. Multivariate Regression Model Results of Effect of Race on Log of Waiting Time, 2003−2017

		Baseline Model			Fixed Effects Model		
			Black	Hispanic		Black	Hispanic
Year	**N**	R^2	**White [Ref]**	**White [Ref]**	R^2	**White [Ref]**	**White [Ref]**
2003	30,134	0.19	0.13** (0.04, 0.22)	0.24*** (0.15, 0.33)	0.36	0.04* (0.003, 0.08)	0.10*** (0.05, 0.15)
2004	27,948	0.15	0.17*** (0.09, 0.24)	0.04 (-0.04, 0.13)	0.31	0.03 (0.00, 0.06)	-0.01 (-0.06, 0.04)
2005	26,271	0.16	0.08 (−0.01, 0.17)	0.19*** (0.12, 0.26)	0.33	−0.01 (−0.05, 0.03)	0.06* (0.01, 0.11)
2006	26,770	0.15	0.19** (0.06, 0.32)	0.18** (0.06, 0.30)	0.35	0.03 (−0.01, 0.08)	0.05 (−0.01, 0.10)
2007	23,496	0.15	0.24*** (0.12, 0.37)	0.14* (0.02, 0.26)	0.33	0.03 (−0.03, 0.09)	0.04 (−0.02, 0.09)
2008	22,844	0.17	0.14** (0.06, 0.21)	0.03 (−0.07, 0.14)	0.33	0.05* (0.01, 0.09)	0.03 (−0.04, 0.10)
2009	26,726	0.09	0.24*** (0.11, 0.37)	0.12 (−0.01, 0.25)	0.32	0.01 (−0.03, 0.06)	0.03 (−0.02, 0.09)
2010	27,581	0.09	0.20** (0.08, 0.32)	0.03 (−0.06, 0.11)	0.30	−0.01 (−0.06, 0.03)	0.03 (−0.03, 0.09)
2011	24,176	0.09	0.16* (0.03, 0.30)	−0.03 (−0.16, 0.10)	0.29	0.04 (−0.01, 0.10)	0.00 (−0.06, 0.07)
2012	21,409	0.08	0.23*** (0.12, 0.34)	0.11 (−0.03, 0.25)	0.38	0.03 (−0.03, 0.09)	0.05 (−0.01, 0.11)
2013	17,820	0.07	0.12 (−0.04, 0.27)	0.02 (−0.11, 0.14)	0.31	−0.01 (−0.08, 0.06)	−0.03 (−0.10, 0.04)
2014	17,408	0.11	−0.14 (−0.43, 0.16)	−0.05 (−0.22, 0.12)	0.48	0.03 (−0.05, 0.12)	0.06 (0.00, 0.12)
2015	15,035	0.08	−0.12 (−0.27, 0.03)	0.01 (−0.20, 0.23)	0.49	−0.05 (−0.13, 0.03)	0.02 (−0.07, 0.10)
2016	14,237	0.08	−0.10 (−0.33, 0.12)	−0.07 (−0.30, 0.17)	0.47	0.01 (−0.05, 0.08)	0.08 (−0.01, 0.16)
2017	12,165	0.05	0.05 (−0.13, 0.24)	0.05 (−0.13, 0.23)	0.39	−0.04 (−0.10, 0.03)	0.07 (−0.01, 0.15)

Notes: Abbreviations: ref, reference. Data were weighted to be nationally representative. Adjusted covariates were age, sex, race/ethnicity, source of payment, primary reason for visit, acuity level, mode of arrival, ED region, ED metro status, season, weekday (dummy), work shift, and resident/intern physician involvement. *Metropolitan status was not reported in 2012 in the NHAMCS public use file.

Whites, 1.8% (95% CI: 0.5% to 3.1%) for Blacks, and by 3.4% (95% CI: 2.3% to 4.4%) for Hispanics. For bedside registration, only Hispanics saw an increase of 1% (95% CI: 0.2% to 1.8%), and with physician-at-triage, only Blacks and Hispanics had annual increases of 2.9% (95% CI: 0.1% to 5.6%) and 3.7% (95% CI: 0.7% to 6.7%), respectively. There were no statistically significant changes in the percent of patients seen at EDs that indicated they used fast track and immediate bedding strategies.

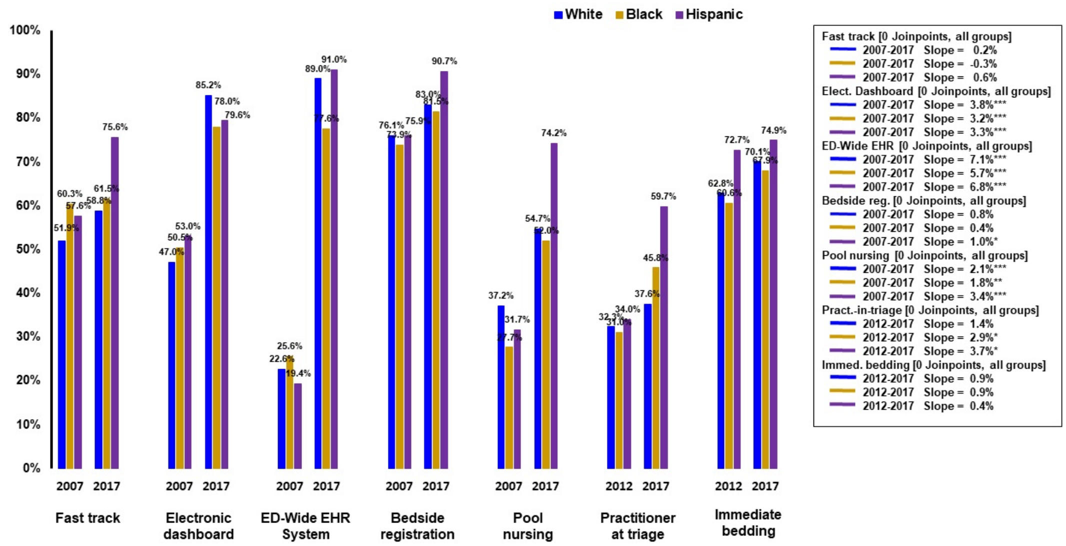

Figure 3.3. Percent of patients of each racial/ethnic group treated at EDs that implemented specific throughput solutions, 2007/2012 versus 2017. Data were weighted to be nationally representative.

Note: Abbreviations: EHR – electronic health record; Elect. -– electronic; reg. – registration; Pract. -– Practitioner; Immed. — immediate. EHR trend analysis does included data from 2007−2009 and 2014−2017 (2010 to 2013 excluded because the exact variable was not reported during that period).

3.5 Discussion

Prolonged waiting in the ED could compromise care quality. Yet, previous studies documented notable differences in how long Blacks and Hispanics waited to see an emergency practitioner compared to Whites. This study provides an update on national trends in ED wait time by analyzing data from 2003 to 2017. Over time, the median wait time decreased to under twenty minutes across the three racial and ethnic groups. Even more, the data indicate that racial and ethnic disparities have disappeared. These improvements occurred despite increased visits across all three groups over time (See Table 2.2). Previous results on the source of the disparities are mixed. The results of this study show that the disparities emanated almost exclusively from variations in how low acuity patients across the three groups waited to be seen at select EDs.

This is excellent news considering the toll that prolonged waiting can have on health outcomes. Evidence suggests that timely evaluation and treatment of patients with conditions such as stroke, heart attack, and pneumonia reduces the likelihood of long-term disability and mortality [102, 103, 104]. For instance, the popular phrase "time is brain" emphasizes the impact of prompt attention to patients with stroke symptoms. The results indicate that the disparities in wait time were concentrated among low acuity patients. Thus, it may be assumed that minorities' health outcomes may not be affected by delays since time-sensitive conditions are attended to promptly in the ED. However, available evidence suggests that more than 50% of ED patients diagnosed with stroke or acute myocardial infarction are initially triaged as low acuity [105, 106]. Such low triage scores have been shown to cause substantial delays in ECG acquisition and reperfusion therapy, which negatively impacts health outcomes [107]. Beyond health outcomes, the opportunity costs of waiting to receive care can be astronomical [108]. Thus, the observed improvement in ED wait times may have positively impacted patient health outcomes and reduced opportunity costs.

The reasons for the decreased wait times - and disparities therein – may be due to recent efforts targeted towards improving quality in U.S. EDs. First, widely publicized calls for attention to the overburdened status of U.S. EDs in the early to mid-2000s prompted several initiatives over the years. For instance, a 2006 IOM report described the status of U.S. EDs as nearing a "breaking point", highlighting challenges that included delays in care delivery [1, 28]. That same year, leaders in emergency medicine held the first of a series of summits to deliberate ED performance, benchmarking, and quality improvement [28]. In 2009, a review report from the Government Accountability Office (GAO) indicated that patients continued to wait longer than recommended [2]. These and similar informative pieces set the stage for a revolution targeted at improving timeliness and quality in U.S. EDs.

Also, with the disparities concentrated among low acuity patients, it may be assumed that increased use of direct-impacting solutions such as physician-in-triage, fast track routes for low acuity patients, and immediate bedding drove the observed declines in ED wait times [96]. However, the results suggest that the contribution of these strategies was minimal. In fact, there were no observed statistically significant changes in the use of fast track and immediate bedding strategies during the study period.

Gains in efficiency resulting from multiple sources including ED staff and culture changes, and effective communication and care coordination in the ED probably improved patient flow and reduced patient wait times [1]. Over time, the use of electronic dashboards and ED-wide electronic health record (EHR) systems increased markedly. One study found that between 2007 and 2010, patients that visited EDs with advanced health information technology waited, on average, six minutes less to see an ED practitioner when compared to patients that visited EDs with only basic systems [6]. A more recent study found that EHR meaningful use improved quality and potentially reduced disparities [109]. The results of this study indicate that ED wait time took downward trends across all three groups only after implementation of the Health Information Technology for Economic and Clinical Health Act (HITECH) Act in 2009, which created incentives for the widespread adoption and use of HIT by health care providers [110].

However, ED wait time disparities remained after HITECH. In adjusted models, the disparities between Blacks and Whites persisted from 2003 to 2012 and that between Hispanics and Whites was almost non-existent. The sudden disappearance of disparities across all groups coincided with the implementation of the CMS pay-for-reporting and public reporting initiatives in 2013 [30]. While the results suggest that disparities have actually disappeared, differences in how patient wait time was measured by EDs for reporting purposes could affect this result [33]. Overall, several recent innovations could reduce ED patient wait times but the data does not capture those variables.

This study has limitations. First, it is possible that NHAMCS data is not fully representative of the U.S. population. The data may also have some inaccuracies, especially because they are abstracted from medical records. Also, NHAMCS data undergo quality control checks to ensure reliability [111]. In particular, starting in 2005, wait time data were manually checked for consistency [39]. The analysis was based on survey data, which makes it difficult to directly attribute the observed improvements in ED wait time to a single intervention. Also, whether the improvement translated into improved health outcomes could not be assessed.

3.6 Conclusion

We found significant decreases in ED wait time across White, Black, and Hispanic racial and ethnic groups, and what appears to be a complete disappearance of disparities between Whites and the two minority groups. This improvement may be associated with recent initiatives targeted towards improving general quality of care in U.S. EDs.

CHAPTER 4

PREDICTING EMERGENCY DEPARTMENT PATIENT DISPOSITION

4.1 Introduction

Previously just a room within a hospital reserved for emergency cases, the emergency department (ED) has grown into an integral unit of any health care system. Between 2006 and 2016, patient visits to the ED in the U.SStates increased from 119.2 to 145.6 million (Table 2.2). As the volume of visits continue to increase, and concerns about crowding remain [112], the need for speed and efficiency in the ED is more important than ever.

Most EDs assign illness severity scores to patients as part of measures to improve resource allocation and overall efficiency. Other recent measures include the creation of separate routes for certain groups of patients (e.g., elderly, pregnant women) and the assignment of a physician to the triage unit to make early patient admission decisions. The latter practice, in particular, has been shown to improve decision-making and patient flow without compromising quality of care [113].

With the widespread adoption of health information technologies used to capture patient data, experts have called for the use of artificial intelligence (AI) to support the efficiency improvement agenda [114, 115]. One aspect of emergency care that could benefit from AI is patient admission decision-making. Delays resulting from failure in communication between inpatient bed management and the ED admissions team have often resulted in the boarding of admitted patients in ED hallways or similar units [116], which can have severe negative effects on quality of care. Predicting patient disposition in advance could improve communication between the two teams, and thus improve efficiency in the ED.

Earlier attempts to predict patient disposition relied on (mostly individual) expert knowledge and experience. While experienced medical personnel (e.g., nurses, physicians, paramedics) achieved high specificity (i.e. successfully identifying patients who will eventually be discharged), their performance on sensitivity (identifying patients who will eventually be admitted to the hospital) was mostly mediocre. For instance, in a prospective cohort study, thirty five experienced physicians at a large urban tertiary care teaching hospital who predicted patient disposition achieved a specificity score of 89.1%, but only a sensitivity score of 51.8% [117].

With the advent of machine learning (ML), attempts have been made to use algorithms to predict patient disposition. The argument in favor of this is that, in general, a patient presents with a symptom or complaint (mostly an initial visit in the case of an ED), and the physician responds by collecting and analyzing all the patient's information at his/her disposal and then making a tentative diagnosis. Since models are better at integrating and synthesizing large data sets from multiple sources, it is argued that, they can integrate data on previous decisions made by multiple physicians involving multiple patients to provide clinical support for improved decision-making [115].

But previous models relied only on structured data such as patient age, sex, vital signs, and other similar structured quantitative variables to predict patient disposition. In some of these instances, the models did not perform better than medical staff partly because of the limited patient information fed to these models [118].

Recent advances in natural language processing (NLP) have made it possible to incorporate narrative text into these models. Several studies have found that using such text alone [119, 120, 121] or combining that with other relevant patient variables collected at triage improves model performance [122, 123, 124, 125].

But these text documented in the expressions of a health practitioner (generally referred to as "free-text") may be biased and limited in detail, and thus fail to adequately explain the variation in patient disposition. Take for instance, an arriving patient with the following reason-for-visit narrative: *"I fell off a horse. Been feeling severe pain in my shoulder for the past two weeks."* An experienced practitioner may simply document *"fall"* or *"pain in (upper) extremity"* as the primary complaint. But the words *horse*, *severe*, and *weeks* emphasize the intensity and duration of the patient's condition, which may explain the variability in patient disposition.

This study makes two major contributions to the literature. First, the use of verbatim patient reason-for-visit narrative text to predict patient disposition is new. Second, previous studies on this subject focused on achieving higher prediction accuracy. This study goes beyond model accuracy to assess model explainability (the extent to which words in a given expression influenced a given prediction).

4.2 Literature Review

The incorporation of text into statistical models is fairly recent. As such, only a few studies have used text data collected at triage to predict patient disposition. Some of these have used only narrative text [119, 120, 121] while others have combined text with other structured data [122, 123, 124, 125].

For instance, Tahayori and colleagues [119] predicted patient disposition using only triage notes. They used Birectional Encoder Representations from Transformers (BERT) model for prediction. The data involved patients that presented to the ED at the St. Vincent's Hospital in Melbourne Australia, between July 2010 and June 2019. The accuracy of the algorithm was 83% and the area under the curve was 0.88.

Sterling and colleagues [120] also used only "free-form" text data documented by the triage nurse to predict patient disposition (admission or discharge). They trained neural network models on the text processed using bag-of-words, paragraph vectors, and topic distributions. They concluded that the best performing model was the one trained on the text processed via the paragraph vector approach, with an AUC of 0.79.

Similarly, Lucini and colleagues [121] used several text mining and machine learning procedures to predict hospitalization. Using text reports (before lab tests) on 11,175 patients at a teaching hospital. They concluded that the best performing model was a support vector machine (SVM) which attained an f1-score of 0.78.

Chen and colleagues [122] combined structured data with physician narratives to predict patient disposition. Using data on 85,775 patient encounters, they found that the incorporation of "the first physician's clinical narratives" improved model performance beyond that achieved using only the structured data.

Zhang and colleagues [123] investigated whether the incorporation of standardized reason for visit text improved model performance. Unlike the other studies, this text was not "free-text", rather it was text from the reason for visit classification for ambulatory care standardized by the National Center for Health Statistics [99]. Using logistic and multi-layer neural network models with principal components analysis (PCA), they found that model performance improved on the inclusion of text data.

Fernandes and colleagues [124] combined patient chief complaint with routinely collected data at triage to build models capable of identifying which patients are at a higher risk of ICU admission. The sample data included patients 18 years or older who visited hospitals in Portugal (253,826 ED visits) and the U.S. (120,649 ED visits). The chief complaint text data for the U.S. dataset was standardized according to the Hierarchical Presenting Problem ontology (HaPPy) [126] while that of the Portuguese hospital was "free-form" text. They concluded that the combination of structured variables (vital signs and other patient-level variables in this case) and patient complaint text data for predicting the disposition of low acuity patients improved model performance.

Roquette and colleagues [125] also used both structured and unstructured data available at triage time to predict pediatric patient disposition. Unstructured text data included the list of previous medications, chief complaint, "free-form" triage notes, imaging, and other tests requested during patient's last visit. They concluded that the addition of textual data significantly improved model performance.

4.3 Methods and Data

This was a retrospective study using data on sample patient in 2005 and 2006. The data were collected through the National Ambulatory Medical Survey (NHAMCS) through which approximately 100 patient visits were sampled from several hospital-based EDs in the U.S. in 2005 and 2006. A total of 69,454 patients were sampled in those two years. Of those, 1,660 did not have either disposition or reason for visit narrative recorded and were excluded from the analysis.

4.3.1 Outcome Variable

The outcome variable was patient disposition. The patient disposition variable has four major classifications: (1) admission to the critical care unit; (2) admission to cardiac catheterization lab (cath lab); (3) admission to hospital; and (4) not admitted (i.e. discharged to go home). All admissions, (1), (2), and (3), were combined into an "admit" class.

4.3.2 Predictor Variable: Verbatim Reason for Visit Narrative

While NHAMCS data collection continues to date, it was only in years 2005 and 2006 that verbatim patient reason(s) for visit ("using the patient's own words if possible") was collected. These verbatim text was subsequently recoded into standardized reason for visit text and codes according to the reason for visit classification developed by the National Center for Health Statistics (NCHS) [99].

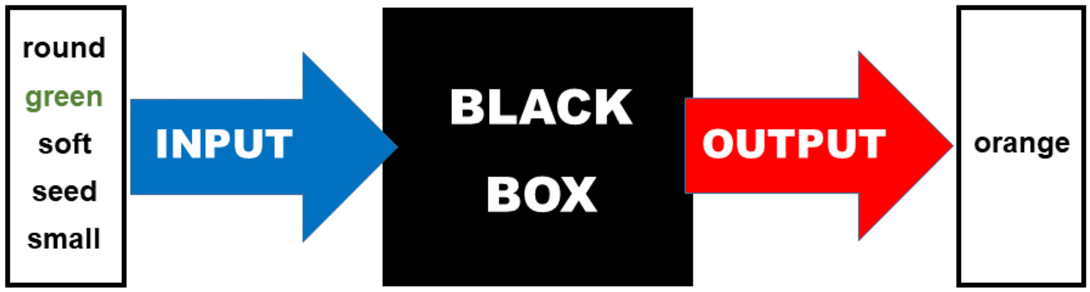

Figure 4.1. Illustration of a black-box model with text input.

4.3.3 Statistical Analysis

Text Pre-Processing. The text were preprocessed by removing numbers and commonly occurring words that are generally not relevant to the prediction task (e.g., prepositions). Word stemming and lemmatization is also performed to transform a word that is used in different forms to its root (e.g., transform "falling" to "fall" by removing the "ing" or "cats" to "cat"). Since a single word may not differentiate patients, the analysis was limited to reason for visit expressed in at least two words (generally, the more words there are, the better). After this process, the sample size reduced to 41,784 observations.

Prediction Models. The raw pre-processed text were fed a deep neural network model for training and prediction. The deep learning model used is illustrated in Figure 4.2. The text vectorization layer takes the string input and converts it to a tensor (a multidimensional array). The embedding layer, which requires that the input data be integer encoded, accepts the transformed tensor as input. Next is a global average pooling layer which returns a fixed-length output vector by averaging the word vectors. The output of this layer is fed to a fully connected (dense) layer (with 16 nodes) with a rectified linear activation function or ReLU. Next to this is a layer with a drop out rate of 50%. This layer is connected to a single output node with a sigmoid activation function that outputs a value between 0 and 1 representing the probability of disposition.

```
Model
Model: "model"
_________________________________________________________________________________________________
Layer (type)                             Output Shape                        Param #
=================================================================================================
input_1 (InputLayer)                     [(None, 1)]                         0
_________________________________________________________________________________________________
text_vectorization (TextVectorization)   (None, 15)                          0
_________________________________________________________________________________________________
embedding (Embedding)                    (None, 15, 16)                      91872
_________________________________________________________________________________________________
global_average_pooling1d_1 (GlobalAveragePo (None, 16)                       0
_________________________________________________________________________________________________
dense_3 (Dense)                          (None, 16)                          272
_________________________________________________________________________________________________
dropout_1 (Dropout)                      (None, 16)                          0
_________________________________________________________________________________________________
dense_2 (Dense)                          (None, 1)                           17
=================================================================================================
Total params: 92,161
Trainable params: 92,161
Non-trainable params: 0
_________________________________________________________________________________________________
```

Figure 4.2. Structure of the deep neural network model.

Baseline Model. The baseline model was a logistic regression model. This model will be compared to the deep learning model.

Model Training, Validation, & Testing A random sample of 70% of the data was used for model training and validation, and the remaining 30% for testing. Since, the data were imbalanced, a higher weight was assigned to the minority class during model training to avoid overfitting.

The batch size and number of epochs were set to 30 and 512, respectively. The learning rate was set to 0.25 and the optimizer was set to adam. The cross entropy loss between the labels and predictions was computed with sparse categorical cross entropy. Model performance was assessed using balanced accuracy, f1-score, area under the curve, sensitivity, and specificity.

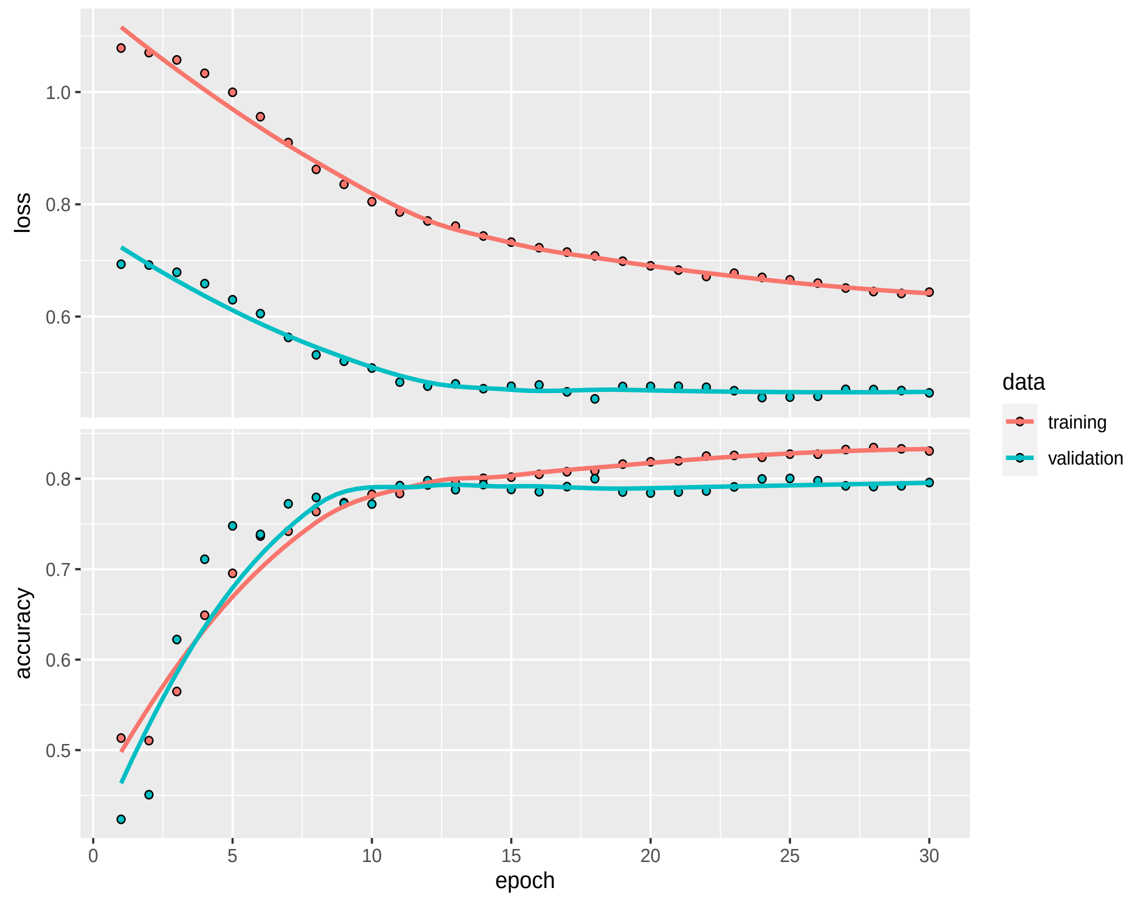

Figure 4.3. Deep neural network model training and validation.

Model Explainability The Local Interpretable Model-agnostic Explanations (LIME) [**?**] was used to assess the influence of words in a given sequence of text for a given prediction. LIME is model-agnostic, which means it can be applied to any model. LIME also focuses on explaining individual instances. This is particularly important in this case because of the need to focus on individual patient conditions - two patients with *headache* may have different diagnoses.

The idea behind LIME is to approximate a complex black-box model with a simple linear model on a local scale. LIME for text data proceeds as follows:
(1) Given an instance or observation, permute it to create replicated data. In the case of text, this permutation is performed by randomly removing words from the original

sequence (*e.g., given an original sequence: patient complains of stomach pain; sample permuted sequence 1: patient of pain); sample permuted sequence 2: patient complains stomach)*; ... (2) next, predict probabilities for the numerous (e.g., 5000) permutations of distorted sequences using the black-box model; (3) estimate the distance between the original instance and the permuted instances (called weights or similarity scores) – calculated as 1 minus the proportion of words that were removed (in the given example $weight = 1-2/5$); (4) select n features from the permuted data that best describe the complex model; (5) fit a simple model to the permuted data (for the n features) that best approximates the complex model. Here, the permuted data is weighted by the similar scores computed in step 3; (6) finally, extract the feature weights from the simple model and these serve as some feature importance to explain the complex model for the given instance. This procedure was implemented using the *lime package* [127] in *R*.

4.3.4 Participants

Of the 67,794 patients, 41,784 had reasons for visits expressed in at least two words. After preprocessing text, the first and third quantiles for text length were 2 and 5, respectively. The median and mean text length were 3 and 3.9, respectively. The maximum text length was 15 words. Approximately 10.7% of the patients were admitted to the hospital.

4.3.5 Model Prediction Performance

Table 4.1 shows the prediction performance on the deep learning model on the full test set as well across the subgroups of symptomic and non-symptomic conditions. Overall, the model achieved a balanced accuracy of 0.79 on the test set. The AUC was 0.82 and f1 score was 0.87. The sensitivity and specificity values were 0.90 and 0.69, respectively. For the symptomatic conditions only, balanced accuracy was 0.80. The AUC was 0.84 and F1 score was 0.87. The sensitivity and specificity values were 0.88 and 0.72, respectively. Across the subgroups of symptoms, balanced accuracy ranged from 0.70 to 0.82, with skin, nails, and hair symptoms having the least balanced accuracy and general symptoms having the highest balanced accuracy. For the non-symptomatic conditions, balanced accuracy was 0.75. The AUC was 0.81 and f1 score was 0.87. The sensitivity and specificity values were 0.88 and 0.63, respectively.

The logistic regression appeared to overfit the data and performance was poor overall (Table 4.2).

Table 4.1. Deep learning Model Performance on Predicting Patient Disposition

Sample	**N**	**Bal. Acc.**	**F1 Score**	**AUC**	**Sensitivity**	**Specificity**
All (test set)	12,536	0.79	0.87	0.82	0.90	0.69
Symptom	9,485	0.80	0.87	0.84	0.88	0.72
General	2,204	0.82	0.85	0.84	0.85	0.78
Psychological and mental Disorders	550	0.71	0.79	0.74	0.78	0.64
Nervous System	751	0.74	0.84	0.80	0.86	0.62
Cardiovascular and Lymphatic System	89	0.66	0.69	0.72	0.64	0.68
Eyes and Ears	357	0.76	0.97	0.81	0.98	0.55
Respiratory System	1,338	0.81	0.86	0.88	0.87	0.75
Digestive System	1,695	0.78	0.81	0.86	0.86	0.70
Genitourinary System	541	0.72	0.88	0.78	0.88	0.57
Skin/Nails/Hair	317	0.70	0.91	0.84	0.90	0.50
Musculoskeletal System	1,643	0.72	0.90	0.80	0.92	0.52
Other (non-symptom)	3,051	0.75	0.87	0.81	0.88	0.63

Note: Model Positive Class: ”0”; Model Negative Class: ”1”.

4.3.6 Interpreting Model Predictions

Figure 4.4 shows the lime algorithm’s performance on a few instances of the test set. Each case or instance shows the predicted label with its probability of occurrence. It also shows the extent to which the simple linear model approximates or explains the complex model for the given case (Explanation Fit). The plot shows how each word explains the linear model in the local region for the case. A word may support (by increasing the probability of a given label) or contradict a given prediction (by decreasing the probability of a given label). For instance, regarding case 1, the model predicts the patient will be discharged with a probability of 0.88. The simple linear approximates the complex model for this instance (Explanation fit = 0.91). The patient's complaint is about the hand and finger hurting as well as numbness in their right knee. The interpretation here is that, the word "numb" reduces the probability of discharge by less than 0.1 while the words "finger" and "hand" increases the probability of discharge by more than 0.1. The words "left", "hurt", "right", and "knee" all increase the probability of discharge by less than 0.1.

A shiny application was also developed based on the deep learning model. This application outputs a predicted label, its probability of occurrence, explainer fit, and explanations for a given sequence of text (in this case reason for visit) by highlighting the words that support or contradict the predicted label. In the randomly entered text shown in Figure 4.5, the model predicts a discharge with a probability of 0.87. The model explainer fits almost perfectly. In this application, the default number of word permutations for approximating the deep learning model is set to 5000 but that can changed. It also provides several word selection strategies: none (i.e. use all features for the explanation); highest weights (use the n features with highest absolute weight in a ridge regression); auto (uses forward selection if $n \leq 6$ and otherwise highest weights); lasso (use the n features least prone to shrinkage based on lasso regularization); forward selection (features are added one by one based on improvement); and tree (a tree is fitted with $log_2(n)$ splits, to use at maximum n features).

Table 4.2. Logistic regression model performance on predicting patient disposition

Sample	N	Bal. Acc.	f1 Score	AUC	Sensitivity	Specificity
All (test set)	12,536	0.67	0.94	0.76	0.90	0.43
Symptom	9,485	0.65	0.94	0.75	0.89	0.40
General	2,204	0.65	0.91	0.74	0.85	0.46
Psychological and mental Disorders	550	0.50	0.86	0.59	0.78	0.21
Nervous System	751	0.69	0.93	0.75	0.88	0.50
Cardiovascular and Lymphatic System	89	0.50	0.99	0.54	1.00	0.00
Eyes and Ears	357	0.49	0.99	0.67	0.98	0.00
Respiratory System	1,338	0.63	0.94	0.79	0.89	0.38
Digestive System	1,695	0.67	0.93	0.61	0.87	0.47
Genitourinary System	541	0.70	0.95	0.75	0.91	0.50
Skin/Nails/Hair	317	0.73	0.98	0.0.70	0.96	0.50
Musculoskeletal System	1,643	0.47	0.97	0.75	0.94	0.00
Other (non-symptom)	3,051	0.71	0.95	0.79	0.92	0.51

Note: Model Positive Class: "0"; Model Negative Class: "1".

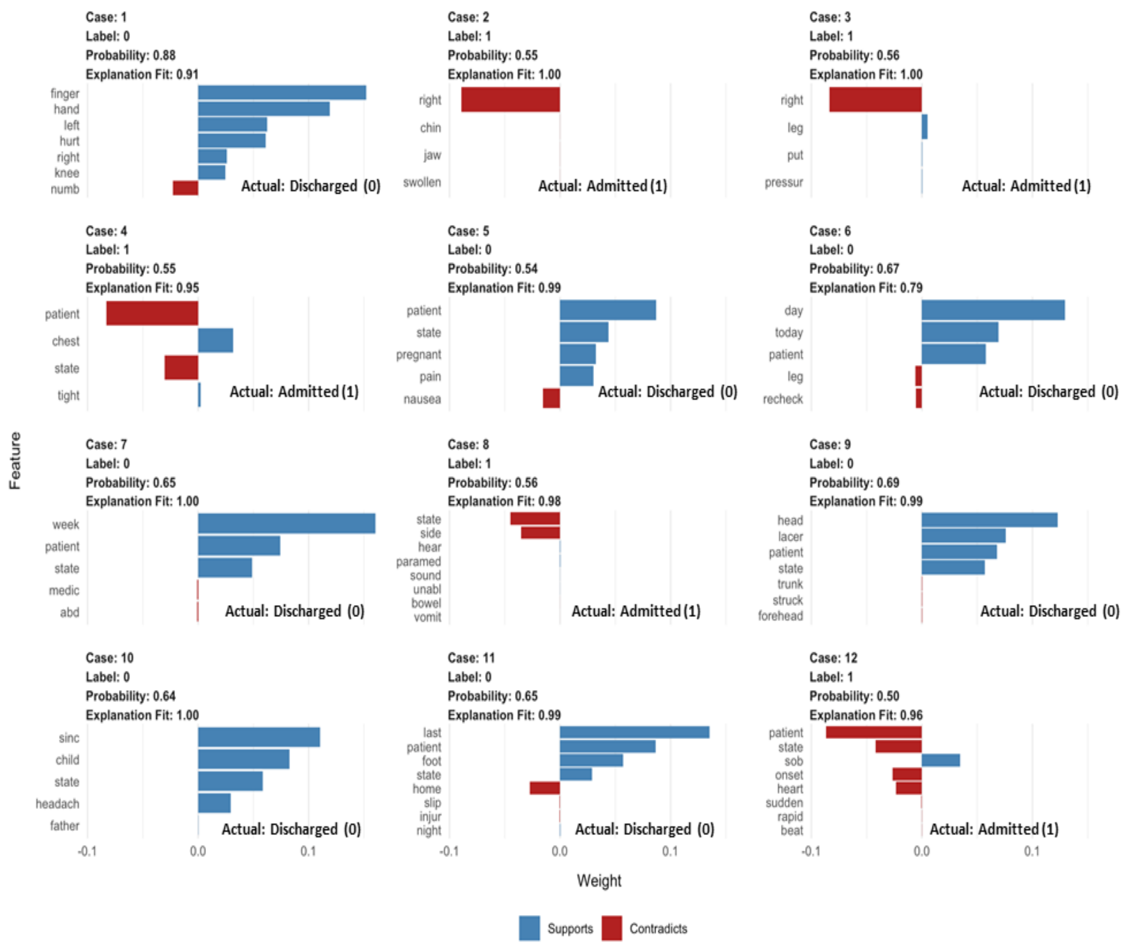

Figure 4.4. Sample model and lime explainer outputs for sample patients.

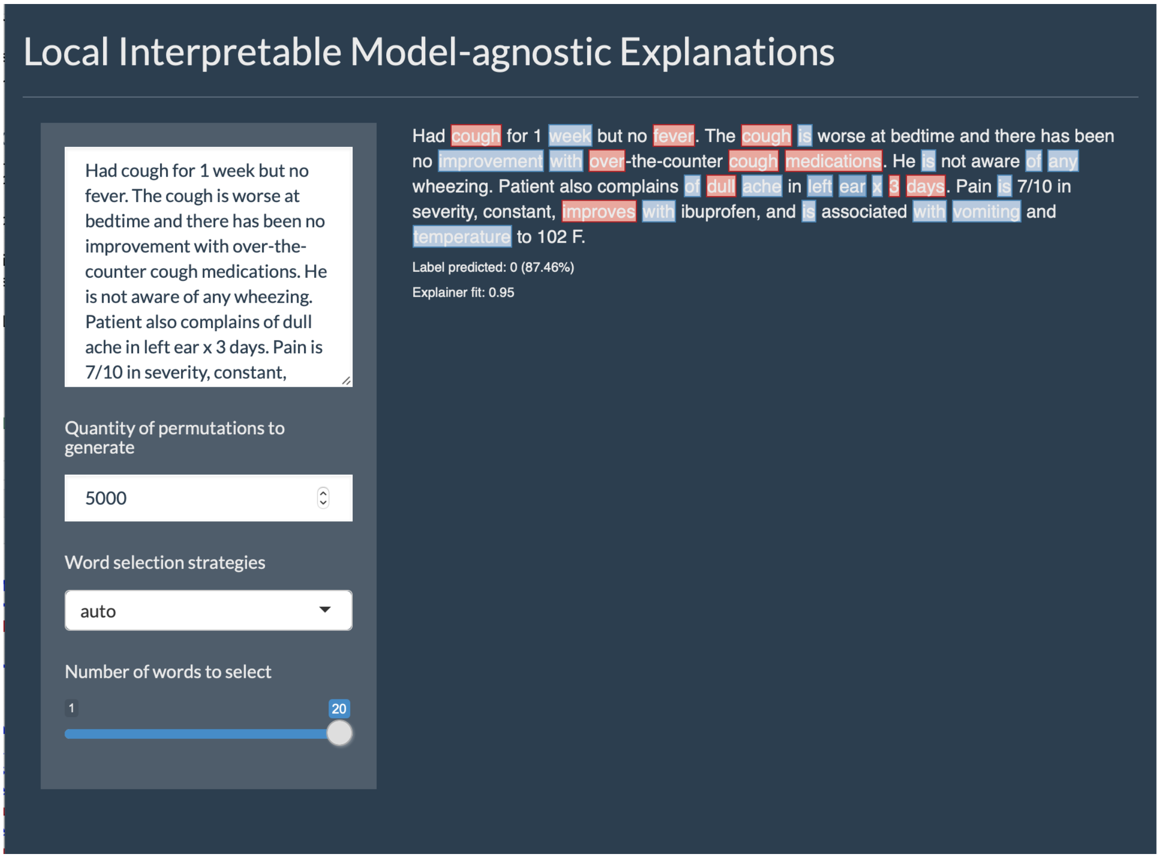

Figure 4.5. LIME explainer.

4.4 Discussion

Delays in care delivery in urgent care settings such as the ED may negatively impact patient health outcomes or even worse, result in patient mortality. The National Academy of Medicine identifies timeliness in care delivery as a major attribute of quality care. In an effort to improve patient flow and reduce crowding in U.S. EDs, the American College of Emergency Physicians encourages the implementation of several proposed throughput strategies [8].

As part of efforts to improve timeliness and efficiency in the ED, the use of advanced health information technology, and more recently, artificial intelligence, have been encouraged. Widespread adoption of health information technology has made it possible to capture a variety of enormous data which has spurred calls for the use of artificial intelligence to improve the quality of care by leveraging such data.

As a result, several recent studies have sought to use clinician notes to make predictions that providers may use to make decisions early in the care process. However, these studies simply showed that triage notes could improve model performance. In this study, an interpretable deep learning model is developed to support early patient disposition decision-making. This model is different in many ways. First, the use of verbatim patient reason for visit narrative text which potentially eliminates or reduces the biases of nurses documenting their observations as opposed to a patient's actual experiences. Also, some predictions are easy to make without any models. For example, patients visiting for a checkup or refill prescriptions are less likely to be admitted. As such, reasons for visits were further stratified into symptom- and non-symptom-based cases. In the former, the patient is most likely making an initial visit and the physician has to make a decision based on limited information. The predictive model achieved high accuracy overall and across several symptom-based conditions.

Also, it is often unclear from previous studies what exactly is influencing a given prediction. While it is useful to achieve high prediction, it is important to provide more actionable information that providers can use to improve decision making. Beyond achieving high performance in predicting patient disposition, an interpretable algorithm was developed to identify the parts of a given sequence of text that influenced a given prediction. Also an application was developed to assess which parts of any given text is predictive of patient disposition.

This study has limitations. First, while it is encouraged to document verbatim patient reason for visit, some patients may not be able to provide a narrative due to their condition. In this case, the next best alternative is probably a narrative from a family member if that is available. If this too is not possible, the narrative will be documented in the words of the triage nurse which may be biased. It is important to note that, even the narrative provided by the patient may not be the most accurate representation of their condition. The incorporation of other variables such as vital signs (temperature, pulse rate, respiratory rate, e.t.c) may also improve model performance.

The *LIME* algorithm used to interpret the predictions is a tool with strengths but also limitations. The simple linear model may not approximate the complex model (in this case a deep learning model) satisfactorily. In such instances, model explainability may be misleading. Another concern closely related to the first is the stability of explanations provided by the model. The explanations of very close points may actually vary greatly in a simulated setting [128].

4.5 Conclusion

In this study, verbatim reason for visit narrative at the triage stage was leveraged to predict patient disposition ahead of time. Importantly, an interpretable model is developed to offer insights as to why a patient would be admitted to the hospital ward. The incorporation of other pertinent variables such as patient demographics, vital signs, and other data available at triage may improve model performance and provide more insights.

APPENDIX A

CHAPTER 2 SUPPLEMENTARY DATA

Extra variables to complement Table 2.2.

Table A.1. Trends in Total Number of Visits, millions (Standard Error), 2006-2016

Characteristic	2006	2016	Absolute Change
Diagnostic screening ordered			
Yes	92.3 (4.6)	38.5 (2.8)	−53.8
No	25.0 (1.9)	105.3 (6.5)	80.3
Work shift			
7:00−15:00	47.6 (2.4)	59.1 (3.6)	11.5
15:00−23:00	52.6 (2.7)	63.5 (4.0)	10.9
23:00−7:00	16.7 (0.9)	20.3 (1.4)	3.6
Week day			
Yes	84.8 (4.2)	105.9 (6.5)	21.1
No	34.4 (1.6)	39.7 (2.5)	5.3
Season			
Winter	32.5 (3.7)	36.2 (6.0)	3.7
Spring	27.6 (3.6)	27.1 (4.6)	−0.5
Summer	30.5 (3.3)	38.2 (5.3)	7.7
Autumn	28.6 (3.9)	44.2 (6.7)	15.6
Resident/Intern involved			
Yes	10.7 (1.4)	11.9 (2.4)	1.2
No	108.4 (5.4)	133.7 (8.6)	25.3

Notes: *Absolute change, calculated as $total\ visits_{2016} - total\ visits_{2006}$.

Extra variables to complement Table 2.3.

Table A.2. Trends in Median (Interquartile Range) Wait Time in Minutes, 2006-2016

Characteristic	2006	2016	Annual Change (95% CI)
Diagnostic screening ordered			
Yes	32 (14, 70)	18 (6, 40)	−2.0 (−2.4 to −1.6)
No	30 (14, 60)	17 (5, 46)	−1.9 (−2.4 to −1.3)
Work shift			
7:00−15:00	31 (14, 65)	17 (6, 42)	−1.8 (−2.1 to −1.4)
15:00−23:00	34 (15, 74)	18 (6, 49)	−2.0 (−2.5 to −1.5)
23:00−7:00	26 (12, 55)	14 (5, 36)	−1.5 (−1.9 to −1.1)
Week day			
Yes	31 (15, 69)	18 (6, 46)	−1.9 (−2.3 to −1.4)
No	32 (14, 65)	16 (5, 41)	−1.8 (−2.2 to −1.4)
Season			
Winter	34 (15, 75)	20 (6, 48)	−1.6 (−2.3 to −0.8)
Spring	33 (15, 72)	15 (6, 39)	−2.3 (−2.9 to −1.8)
Summer	29 (14, 59)	18 (6, 45)	−1.8 (−2.3 to −1.2)
Autumn	30 (14, 66)	16 (5, 45)	−1.7 (−2.4 to −0.9)
Resident/Intern involved			
Yes	36 (15, 81)	19 (6, 46)	−2.7 (−3.8 to −1.7)
No	31 (14, 66)	17 (6, 45)	−1.8 (−2.2 to −1.3)

Extra variables to complement Table 2.4.

Table A.3. Trends in Median (Interquartile Range) Length of Visit in Minutes, 2006-2015

Characteristic	**2006**	**2015**	**Annual Change (95% CI)**
Diagnostic screening ordered			
Yes	169 (103, 266)	185 (119, 282)	0.14 (−1.4 to 1.6)
No	82 (50, 139)	86 (52, 143)	−0.3 (−1.7 to 1.1)
Work shift			
7:00−15:00	151 (87, 246)	156 (92, 250)	−0.3 (−1.7 to 1.1)
15:00−23:00	145 (84, 236)	153 (90, 244)	0.0 (−1.5 to 1.5)
23:00−7:00	151 (81, 264)	150 (89, 257)	−0.5 (−2.4 to 1.4)
Week day			
Yes	151 (87, 249)	158 (94, 256)	−0.2 (−1.8 to 1.4)
No	140 (80, 231)	143 (84, 228)	0.0 (−2.3 to 1.5)
Season			
Winter	157 (92, 254)	153 (94, 246)	−0.6 (−2.5 to 1.3)
Spring	160 (90, 263)	152 (83, 243)	−1.2 (−3.5 to 1.2)
Summer	138 (80, 226)	149 (89, 238)	0.2 (−1.8 to 2.1)
Autumn	141 (79, 231)	167 (100, 275)	0.8 (−1.2 to 2.8)
Resident/Intern involved			
Yes	210 (123, 340)	225 (138, 366)	−1.0 (−6.1 to 4.1)
No	144 (82, 235)	150 (89, 240)	0.0 (−1.4 to 1.4)

Distributions of Analytical and Missing Samples for Boarding Time Outcome Measure

Table A.4. Distributions of Analytical and Missing Samples for Boarding Time Outcome Measure

Characteristic	**Missing sample (N = 6,257)**	**Analytical sample (N = 12,491)**
Age group, N (%)		
0−17	397 (6.5)	837 (5.8)
18−24	366 (4.4)	620 (4.0)
25−44	1,202 (17.1)	2,355 (16.0)
45−64	1,782 (29.3)	3,859 (29.9)
65+	2,510 (42.7)	5,570 (44.3)
Sex, N (%)		
Female	3,359 (54.0)	7,0097 (53.0)
Male	2,898 (46.0)	6,144 (47.0)
Race/Ethnicity, N (%)		
White	4,817 (78.4)	9,991 (77.6)
Black	1,209 (18.0)	2,515 (18.4)
Other	231 (3.5)	735 (4.0)
Source of Payment, N (%)		
Private insurance	1,397 (22.3)	3,125 (24.0)
Medicare	2,584 (44.1)	5,889 (46.5)
Medicaid/CHIP	1,234 (17.0)	2,378 (16.0)
Other	1,042 (16.6)	1,849 (13.5)
Metropolis, N (%)		
Yes	5,655 (87.7)	11,894 (88.3)
No	602 (12.3)	1,347 (11.7)

Annual proportion of outliers stratified by outcome: wait time, length of visit, and boarding time

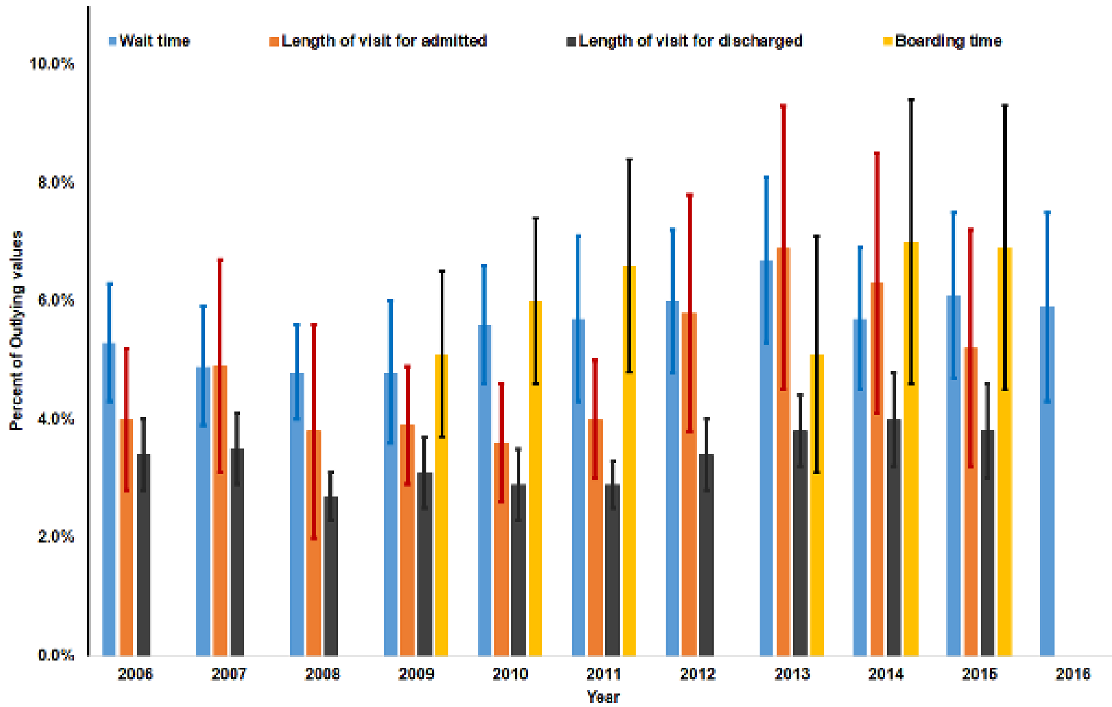

Figure A.1. Annual percent of outliers for wait time measure.

APPENDIX B

CHAPTER 3 SUPPLEMENTARY DATA

Distribution of Missing Data for Wait Time Outcome.

Table B.1. Distribution of Missing Wait Time Sample Stratified by Race/Ethnic Group, 2003-2017

Characteristic	**All**	**White**	**Black**	**Hispanic**
N (%)	73,692 (100)	45,136 (61.2)	16,413 (22.3)	12,143 (16.5)
Age group (years), n (%)				
0−17	17,519 (23.8)	9,371 (20.8)	4,066 (24.8)	4,082 (33.6)
18−44	31,427 (42.6)	18,558 (41.1)	7,689 (46.8)	5,180 (42.7)
45−64	15,280 (20.7)	9,787 (21.7)	3,471 (21.1)	2,022 (16.7)
65+	9,466 (12.8)	7,420 (16.4)	1,187 (7.2)	859 (7.1)
Sex, n (%)				
Female	39,502 (53.6)	24,043 (53.3)	9,024 (55.0)	6,435 (53.0)
Male	34,190 (46.4)	21,093 (46.7)	7,389 (45.0)	5,708 (47.0)
Payment Source, n (%)				
Private insurance	22,171 (32.6)	16,217 (38.5)	3,369 (22.7)	2,585 (23.3)
Medicare	10,343 (15.2)	7,920 (18.8)	1,576 (10.6)	847 (7.6)
Medicaid/CHIP	19,103 (28.1)	8,856 (21.0)	5,815 (39.2)	4,432 (40.0)
Self-Pay	10,635 (15.6)	5,778 (13.7)	2,802 (18.9)	2,055 (18.5)
Other	5,833 (8.6)	3,401 (8.1)	1,258 (8.5)	1,174 (10.6)
Region, n (%)				
Northeast	19,328 (26.2)	12,461 (27.6)	3,576 (21.8)	3,291 (27.1)
Midwest	15,696 (21.3)	10,844 (24.0)	3,554 (21.7)	1,298 (10.7)
South	22,761 (30.9)	12,175 (27.0)	7,432 (45.3)	3,154 (26.0)
West	15,907 (21.6)	9,656 (21.4)	1,851 (11.3)	4,400 (3.8)
Metropolis*, n (%)				
Yes	59,472 (86.1)	34,083 (80.2)	14,518 (95.2)	10,871 (96.2)
No	9,588 (13.9)	8,428 (19.8)	725 (4.8)	435 (3.8)
Acuity level, n (%)				
Emergent	6,962 (9.4)	4,742 (10.5)	1,277 (7.8)	943 (7.8)
Urgent	20,448 (27.7)	12,992 (28.8)	4,293 (26.2)	3,163 (26.0)
Semi-urgent	16,539 (22.4)	10,002 (22.2)	3,911 (23.8)	2,626 (21.6)
Non-urgent	7,345 (10.0)	4,231 (9.4)	1,913 (11.7)	1,201 (9.9)
No triage/Unknown	22,398 (30.4)	13,169 (29.2)	5,019 (30.6)	4,210 (34.7)

Notes: Subgroups may not sum to total N due to missing/unknown responses or sum to 100% due to rounding. *Metropolitan status was not reported in 2012 in the NHAMCS public use file.

REFERENCES

[1] Institute of Medicine. Crossing the quality chasm: A new health system for the 21^{st} century. 2001. Washington, DC: The National Academies Press. `https://doi.org/10.17226/10027`.

[2] Marcia Crosse. Hospital emergency departments: crowding continues to occur, and some patients wait longer than recommended time frames. Darby, Pennsylvania. Diane Publishing, 2010.

[3] American College of Physicians. Racial and ethnic disparities in health care, updated 2010. Philadelphia: American college of Physicians; 2010: Policy Paper. (available from American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106.).

[4] Alan Nelson. Unequal treatment: confronting racial and ethnic disparities in health care. Journal of the National Medical Association, 94(8):666, 2002.

[5] Alan M Garber and Jonathan Skinner. Is american health care uniquely inefficient? Journal of Economic Perspectives, 22(4):27–50, 2008.

[6] Frederic W Selck and Sandra L Decker. Health information technology adoption in the emergency department. Health Services Research, 51(1):32–47, 2016.

[7] Kristy Gonzalez Morganti, Sebastian Bauhoff, Janice C Blanchard, Mahshid Abir, Neema Iyer, Alexandria Smith, Joseph V Vesely, Edward N Okeke, and Arthur L Kellermann. The evolving role of emergency departments in the United States. Rand Health Quarterly, 3(2), 2013.

[8] American College of Emergency Physicians. Emergency Department Crowding: High-Impact Solutions. ACEP White Paper, 2009.

[9] Institute of Medicine (US) Committee on the National Quality Report on Health Care Delivery. Washington (DC): National Academies Press (US). Envisioning the national health care quality report. 2001.

[10] Houser A. Choula R. Feinberg L, Reinhard SC. Valuing the invaluable: 2011 update, the growing contributions and costs of family caregiving. Washington, DC: AARP Public Policy Institute, 32:2011, 2011.

[11] James S House, Karl R Landis, and Debra Umberson. Social relationships and health. Science, 241(4865):540–545, 1988.

[12] David M, Sullivan A.F, Cleary Paul, Rao S, Gordon J.A, Kaushal R., et al. The safety of emergency care systems: results of a survey of clinicians in 65 us emergency departments. Annals of Emergency Medicine, 53(6):715–723, 2009.

[13] Institute of Medicine. Hospital-Based Emergency Care: At the Breaking Point. Washington, DC: National Academies Press, 2006.

[14] Karis L Tekwani, Yaniv Kerem, Chintan D Mistry, Brian M Sayger, and Erik B Kulstad. Emergency department crowding is associated with reduced satisfaction scores in patients discharged from the emergency department. Western Journal of Emergency Medicine, 14(1):11, 2013.

[15] Jesse M Pines, Sanjay Iyer, Maureen Disbot, Judd E Hollander, Frances S Shofer, and Elizabeth M Datner. The effect of emergency department crowding on patient satisfaction for admitted patients. Academic Emergency Medicine, 15(9):825–831, 2008.

[16] Benjamin C Sun, James Adams, E John Orav, Donald W Rucker, Troyen A Brennan, and Helen R Burstin. Determinants of patient satisfaction and willingness to return with emergency care. Annals of Emergency Medicine, 35(5):426–434, 2000.

[17] Tara N Cassidy-Smith, Brigitte M Baumann, and Edwin D Boudreaux. The disconfirmation paradigm: throughput times and emergency department patient satisfaction. The Journal of Emergency Medicine, 32(1):7–13, 2007.

[18] K John McConnell, Christopher F Richards, Mohamud Daya, Cody C Weathers, and Robert A Lowe. Ambulance diversion and lost hospital revenues. Annals of Emergency Medicine, 48(6):702–710, 2006.

[19] Jesse M Pines, Robert J Batt, Joshua A Hilton, and Christian Terwiesch. The financial consequences of lost demand and reducing boarding in hospital emergency departments. Annals of Emergency Medicine, 58(4):331–340, 2011.

[20] Lindsey Woodworth and James F Holmes. Just a minute: the effect of emergency department wait time on the cost of care. Economic Inquiry, 58(2):698–716, 2020.

[21] Paul Krochmal and Tamrah A Riley. Increased health care costs associated with ED overcrowding. The American Journal of Emergency Medicine, 12(3):265–266, 1994.

[22] Qing Huang, Amardeep Thind, Jonathan F Dreyer, and Gregory S Zaric. The impact of delays to admission from the emergency department on inpatient outcomes. BMC Emergency Medicine, 10(1):1–6, 2010.

[23] Jesse M Pines, Charles V Pollack Jr, Deborah B Diercks, Anna Marie Chang, Frances S Shofer, and Judd E Hollander. The association between emergency department crowding and adverse cardiovascular outcomes in patients with chest pain. Academic Emergency Medicine, 16(7):617–625, 2009.

[24] Jesse M Pines and Judd E Hollander. Emergency department crowding is associated with poor care for patients with severe pain. Annals of Emergency Medicine, 51(1):1–5, 2008.

[25] Ula Hwang, Lynne Richardson, Elayne Livote, Ben Harris, Natasha Spencer, and R Sean Morrison. Emergency department crowding and decreased quality of pain care. Academic Emergency Medicine, 15(12):1248–1255, 2008.

[26] Astrid Guttmann, Michael J Schull, Marian J Vermeulen, and Therese A Stukel. Association between waiting times and short term mortality and hospital admission after departure from emergency department: population based cohort study from ontario, canada. British Medical Journal, 342, 2011.

[27] Peter C Sprivulis, Julie-Ann Da Silva, Ian G Jacobs, George A Jelinek, and Amanda RL Frazer. The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments. Medical Journal of Australia, 184(5):208–212, 2006.

[28] Shari Welch, James Augustine, Carlos A Camargo Jr, and Charles Reese. Emergency department performance measures and benchmarking summit. Academic Emergency Medicine, 13(10):1074–1080, 2006.

[29] Shari J Welch, Brent R Asplin, Suzanne Stone-Griffith, Steven J Davidson, James Augustine, Jeremiah Schuur, and Emergency Department Benchmarking Alliance. Emergency department operational metrics, measures and definitions: results of the second performance measures and benchmarking summit. Annals of Emergency Medicine, 58(1):33–40, 2011.

[30] Jennifer L Wiler, Shari Welch, Jesse Pines, Jeremiah Schuur, Nick Jouriles, and Suzanne Stone-Griffith. Emergency department performance measures updates: proceedings of the 2014 emergency department benchmarking alliance consensus summit. Academic Emergency Medicine, 22(5):542–553, 2015.

[31] Medicare Program. Hospital Inpatient Prospective Payment System. Federal Register 2012 IPPS Final Rule, 2011.

[32] Centers for Medicare Medicaid Services. Hospital compare website. Available from `https://www.medicare.gov/hospitalcompare/search.html`, Accessed September 30, 2019.

[33] Rui P, Kang K, and Ashman JJ. National hospital ambulatory medical care survey: 2016 center for disease control and prevention (cdc) emergency department summary tables. Available at `https://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency`. Accessed September 20, 2019.

[34] Sayeh Nikpay, Seth Freedman, Helen Levy, and Tom Buchmueller. Effect of the Affordable Care Act Medicaid expansion on emergency department

visits: evidence from state-level emergency department databases. Annals of Emergency Medicine, 70(2):215–225, 2017.

[35] Mark McClelland, Brent Asplin, Stephen K Epstein, Keith Eric Kocher, Randy Pilgrim, Jesse Pines, Elaine Judith Rabin, and Niels Kumar Rathlev. The Affordable Care Act and emergency care. American Journal of Public Health, 104(10):e8–e10, 2014.

[36] Stephen R Pitts, Jesse M Pines, Michael T Handrigan, and Arthur L Kellermann. National trends in emergency department occupancy, 2001 to 2008: effect of inpatient admissions versus emergency department practice intensity. Annals of Emergency Medicine, 60(6):679–686, 2012.

[37] Stephen R Pitts, Frances L Vaughns, Marc A Gautreau, Matthew W Cogdell, and Zachary Meisel. A cross-sectional study of emergency department boarding practices in the United States. Academic Emergency Medicine, 21(5):497–503, 2014.

[38] Julius C Pham, George K Ho, Peter M Hill, Melissa L McCarthy, and Peter J Pronovost. National study of patient, visit, and hospital characteristics associated with leaving an emergency department without being seen: predicting lwbs. Academic Emergency Medicine, 16(10):949–955, 2009.

[39] Leora I Horwitz and Elizabeth H Bradley. Percentage of us emergency department patients seen within the recommended triage time: 1997 to 2006. Archives of Internal Medicine, 169(20):1857–1865, 2009.

[40] Leora I Horwitz, Jeremy Green, and Elizabeth H Bradley. US emergency department performance on wait time and length of visit. Annals of Emergency Medicine, 55(2):133–141, 2010.

[41] Andrew P Wilper, Steffie Woolhandler, Karen E Lasser, Danny McCormick, Sarah L Cutrona, David H Bor, and David U Himmelstein. Waits to see an emergency department physician: US trends and predictors, 1997–2004: Overcrowding by patients unable to obtain nonurgent care elsewhere might be worsening care for those with truly urgent conditions. Health Affairs, 27(Suppl1):w84–w95, 2008.

[42] Steven J Weiss, Amy A Ernst, Robert Derlet, Richard King, Aaron Bair, and Todd G Nick. Relationship between the national ed overcrowding scale and the number of patients who leave without being seen in an academic ed. The American Journal of Emergency Medicine, 23(3):288–294, 2005.

[43] Christopher MB Fernandes, Paula Tanabe, Nicki Gilboy, Loren A Johnson, Rebecca S McNair, Alexander M Rosenau, Peter Sawchuk, David A Thompson, Debbie A Travers, Nancy Bonalumi, et al. Five-level triage: a report from the acep/ena five-level triage task force. Journal of Emergency Nursing, 31(1):39–50, 2005.

[44] NCHS. 2006 NHAMCS Micro-Data File Documentation. Available from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/, Accessed 10/2/2019.

[45] Geoff Hughes. Four hour target for eds: the uk experience. Emergency Medicine Australasia, 22(5):368–373, 2010.

[46] New Zealand Ministry of Health. Health targets: Shorter stays in emergency departments. Available from http://www.health.govt.nz/new-zealand-health-system/health-targets/about-health-targets/health-targets-shorter-stays-emergency-departments, Accessed 09/14/2019.

[47] Expert Panel. Review of elective surgery and emergency access targets under the national partnership agreement on improving public hospital services: supplementary annexure: New south wales—clinical initiative nurses in emergency departments. New South Wales—Clinical Initiative Nurses in Emergency Departments, 2011.

[48] Andrew Affleck, Paul Parks, Alan Drummond, Brian H Rowe, and Howard J Ovens. Emergency department overcrowding and access block. Canadian Journal of Emergency Medicine, 15(6):359–370, 2013.

[49] National Center for Health Statistics. NHAMCS Micro-Data File Documentation. Available from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS, Accessed 10/2/2019.

[50] Timothy P Johnson and Joseph S Wislar. Response rates and nonresponse errors in surveys. Journal of American Medical Association, 307(17):1805–1806, 2012.

[51] Michael Frigge, David C Hoaglin, and Boris Iglewicz. Some implementations of the boxplot. The American Statistician, 43(1):50–54, 1989.

[52] David C Hoaglin, Boris Iglewicz, and John W Tukey. Performance of some resistant rules for outlier labeling. Journal of the American Statistical Association, 81(396):991–999, 1986.

[53] Thomas Lumley. Analysis of complex survey samples. Journal of Statistical Software, 9(1):1–19, 2004.

[54] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. ISBN 3-900051-07-0.

[55] United States Census Bureau. Current Population Survey (CPS). Available at https://www.census.gov/cps/data/cpstablecreator.html, Accessed December 15, 2019.

[56] Richard C Wuerz, Leslie W Milne, David R Eitel, Debbie Travers, and Nicki Gilboy. Reliability and validity of a new five-level triage instrument. Academic Emergency Medicine, 7(3):236–242, 2000.

[57] U.S. Government Accountability Office. Hospital Emergency Departments. Crowded Conditions Vary Among Hospitals and Communities, 2003. Washington DC.

[58] Institute of Medicine Committee on the Future of Emergency Care in the US Health System. The future of emergency care in the United States health system. Annals of Emergency Medicine, 48(2):115–120, 2006.

[59] Suzanne Mason, Ellen J Weber, Joanne Coster, Jennifer Freeman, and Thomas Locker. Time patients spend in the emergency department: England's 4-hour rule—a case of hitting the target but missing the point? Annals of Emergency Medicine, 59(5):341–349, 2012.

[60] Khic-Houy Prang, Rachel Canaway, Marie Bismark, David Dunt, and Margaret Kelaher. The impact of australian healthcare reforms on emergency department time-based process outcomes: An interrupted time series study. PloS one, 13(12):e0209043, 2018.

[61] Roberto Forero, Nicola Man, Shizar Nahidi, Gerard Fitzgerald, Daniel Fatovich, Mohammed Mohsin, Hanh Ngo, et al. When a health policy cuts both ways: impact of the national emergency access target policy on staff and emergency department performance. Emergency Medicine Australasia, 32(2):228–239, 2020.

[62] Hughes Geoff. The four hour target; problems ahead. Emergency Medical Journal., 2006. doi:10.1136/emj.2005.031948.

[63] Ian Higginson. Life in the 4-hour lane: Get ready for a bumpy ride! Medicus, 50(6):12, 2010.

[64] Judy Lowthian and Peter Cameron. Improving timeliness while improving the quality of emergency department care. emergency medicine australasia. 2012. 24 (3); 219-221.

[65] National Archives Office of the Federal Register and Records Administration. Rules and Regulations. Federal Register, 2017. Available at `https://www.govinfo.gov/app/details/FR-2017-12-14`, Accessed 2/19/2020.

[66] Manya F Newton, Carla C Keirns, Rebecca Cunningham, Rodney A Hayward, and Rachel Stanley. Uninsured adults presenting to US emergency departments: assumptions vs data. Journal of American Medical Association, 300(16):1914–1924, 2008.

[67] Malcolm B Doupe, Dan Chateau, Alecs Chochinov, Ellen Weber, Jennifer E Enns, Shelley Derksen, Joykrishna Sarkar, Michael Schull, Ricardo Lobato de Faria,

Alan Katz, et al. Comparing the effect of throughput and output factors on emergency department crowding: a retrospective observational cohort study. Annals of Emergency Medicine, 72(4):410–419, 2018.

[68] Jesse M Pines, Joshua A Hilton, Ellen J Weber, Annechien J Alkemade, Hasan Al Shabanah, Philip D Anderson, et al. International perspectives on emergency department crowding. Academic Emergency Medicine, 18(12):1358–1370, 2011.

[69] James Choi and Ilene Claudius. Decrease in emergency department length of stay as a result of triage pulse oximetry. Pediatric Emergency Care, 22(6):412–414, 2006.

[70] Marc H Gorelick, Kenneth Yen, and Hyun J Yun. The effect of in-room registration on emergency department length of stay. Annals of Emergency Medicine, 45(2):128–133, 2005.

[71] Jon Mark Hirshon, Thomas D Kirsch, William K Mysko, and Gabor D Kelen. Effect of rotational patient assignment on emergency department length of stay. The Journal of Emergency Medicine, 14(6):763–768, 1996.

[72] Adam J Singer, Joshua Ardise, Janet Gulla, and Julie Cangro. Point-of-care testing reduces length of stay in emergency department chest pain patients. Annals of Emergency Medicine, 45(6):587–591, 2005.

[73] Lorne L Holland, Linda L Smith, and Kenneth E Blick. Reducing laboratory turnaround time outliers can reduce emergency department patient length of stay: an 11-hospital study. American Journal of Clinical Pathology, 124(5):672–674, 2005.

[74] Elizabeth Lee-Lewandrowski, Daniel Corboy, Kent Lewandrowski, Julia Sinclair, Steven McDermot, and Theodore I Benzer. Implementation of a point-of-care satellite laboratory in the emergency department of an academic medical center: impact on test turnaround time and patient emergency department length of stay. Archives of Pathology & Laboratory Medicine, 127(4):456–460, 2003.

[75] Peter K Lindenauer, Denise Remus, Sheila Roman, Michael B Rothberg, Evan M Benjamin, Allen Ma, and Dale W Bratzler. Public reporting and pay for performance in hospital quality improvement. New England Journal of Medicine, 356(5):486–496, 2007.

[76] American College of Emergency Physicians. Publishing wait times for emergency department care: An information paper. Report, American College of Emergency Physicians, Baltimore, 2012.

[77] Jennifer L Wiler, Christopher Gentle, James M Halfpenny, Alan Heins, Abhi Mehrotra, Michael G Mikhail, and Diana Fite. Optimizing emergency

department front-end operations. Annals of Emergency Medicine, 55(2):142–160, 2010.

[78] JA Porter, K Frye, R Rodriguez, M Velez Arias, and G VanDenburg. Impact of implementing a direct bedding process on emergency department throughput and leaving without being seen rates. Annals of Emergency Medicine, 72(4):S11–S12, 2018.

[79] YF Choi, TW Wong, and CC Lau. Triage rapid initial assessment by doctor (triad) improves waiting time and processing time of the emergency department. Emergency Medicine Journal, 23(4):262–265, 2006.

[80] Elizabeth H Bradley, Jeph Herrin, Yongfei Wang, Robert L McNamara, Tashonna R Webster, David J Magid, Martha Blaney, et al. Racial and ethnic differences in time to acute reperfusion therapy for patients hospitalized with myocardial infarction. Journal of American Medical Association, 292(13):1563–1572, 2004.

[81] Lenny López, Andrew P Wilper, Marina C Cervantes, Joseph R Betancourt, and Alexander R Green. Racial and sex differences in emergency department triage assessment and test ordering for chest pain, 1997–2006. Academic Emergency Medicine, 17(8):801–808, 2010.

[82] Nancy Sonnenfeld, Stephen R Pitts, Susan M Schappert, and Sandra L Decker. Emergency department volume and racial and ethnic differences in waiting times in the United States. Medical Care, pages 335–341, 2012.

[83] JM Pines, JS Schwartz, and JE Hollander. Racial disparities in ed wait time prior to physician evaluation and boarding times in the United States. Annals of Emergency Medicine, 3(50):S3, 2007.

[84] Andrew P Wilper, Steffie Woolhandler, Karen E Lasser, Danny McCormick, Sarah L Cutrona, David H Bor, and David U Himmelstein. Waits to see an emergency department physician: US trends and predictors, 1997–2004: Overcrowding by patients unable to obtain nonurgent care elsewhere might be worsening care for those with truly urgent conditions. Health Affairs, 27(Suppl1):w84–w95, 2008.

[85] Kristin N Ray, Amalavoyal V Chari, John Engberg, Marnie Bertolet, and Ateev Mehrotra. Disparities in time spent seeking medical care in the United States. JAMA Internal Medicine, 175(12):1983–1986, 2015.

[86] Carlton Haywood Jr, Paula Tanabe, Rakhi Naik, Mary Catherine Beach, and Sophie Lanzkron. The impact of race and disease on sickle cell patient wait times in the emergency department. The American Journal of Emergency Medicine, 31(4):651–656, 2013.

[87] Christine Y Park, Mary Alice Lee, and Andrew J Epstein. Variation in emergency department wait times for children by race/ethnicity and payment source. Health Services Research, 44(6):2022–2039, 2009.

[88] Catherine A James, Florence T Bourgeois, and Michael W Shannon. Association of race/ethnicity with emergency department wait times. Pediatrics, 115(3):e310–e315, 2005.

[89] Xingyu Zhang, Maria Carabello, Tyler Hill, Kevin He, Christopher R Friese, and Prashant Mahajan. Racial and ethnic disparities in emergency department care and health outcomes among children in the United States. Frontiers in Pediatrics, 7:525, 2019.

[90] William P Qiao, Emilie S Powell, Mark P Witte, and Martin R Zelder. Relationship between racial disparities in ED wait times and illness severity. The American Journal of Emergency Medicine, 34(1):10–15, 2016.

[91] Sudeep J Karve, Rajesh Balkrishnan, Yousef M Mohammad, and Deborah A Levine. Racial/ethnic disparities in emergency department waiting time for stroke patients in the United States. Journal of Stroke and Cerebrovascular Diseases, 20(1):30–40, 2011.

[92] Amal N Trivedi, Wato Nsa, Leslie RM Hausmann, Jonathan S Lee, Allen Ma, Dale W Bratzler, et al. Quality and equity of care in US hospitals. New England Journal of Medicine, 371(24):2298–2308, 2014.

[93] Amal N Trivedi, Alan M Zaslavsky, Eric C Schneider, and John Z Ayanian. Trends in the quality of care and racial disparities in medicare managed care. New England Journal of Medicine, 353(7):692–700, 2005.

[94] Ashwini R Sehgal. Impact of quality improvement efforts on race and sex disparities in hemodialysis. Journal of American Medical Association, 289(8):996–1000, 2003.

[95] Kevin Fiscella, Peter Franks, Marthe R Gold, and Carolyn M Clancy. Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. Journal of American Medical Association, 283(19):2579–2584, 2000.

[96] ACEP Boarding Task Force. Emergency department crowding: high-impact solutions. American College of Emergency Physicians. Working Paper, 2008.

[97] Centers for Medicare and Medicaid Services. Hospital inpatient quality reporting program. `https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html`, 2008.

[98] Patient Protection and Affordable Care Act. Patient protection and Affordable Care Act. Public law, 111(48):759–762, 2010.

[99] Don Schneider, Linda Appleton, and Thoma McLemore. A reason for visit classification for ambulatory care. National Center for Health Statistics. Vital Health Statistics, 2(78), 1979.

[100] National Cancer Institute. Joinpoint Regression Program, Version 4.8.0.1 − April 2020; Statistical Methodology and Applications Branch, Surveillance Research Program, 2020.

[101] Deborah D Ingram, Donald J Malec, Diane M Makuc, Deanna Kruszon-Moran, Renee M Gindi, Michael Albert, Vladislav Beresovsky, Brady E Hamilton, Julia Holmes, Jeannine Schiller, et al. National center for health statistics guidelines for analysis of trends. Vital and health statistics. Series 2, Data Evaluation and Methods Research, (179):1–71, 2018.

[102] Mònica Millán, Laura Dorado, and Antoni Dávalos. Fibrinolytic therapy in acute stroke. Current Cardiology Reviews, 6(3):218–226, 2010.

[103] Peter M Houck, Dale W Bratzler, Wato Nsa, Allen Ma, and John G Bartlett. Timing of antibiotic administration and outcomes for medicare patients hospitalized with community-acquired pneumonia. Archives of Internal Medicine, 164(6):637–644, 2004.

[104] Fibrinolytic Therapy Trialists et al. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. The Lancet, 343(8893):311–322, 1994.

[105] Clare L Atzema, Peter C Austin, Jack V Tu, and Michael J Schull. Emergency department triage of acute myocardial infarction patients and the effect on outcomes. Annals of Emergency Medicine, 53(6):736–745, 2009.

[106] Julie Considine and Bree McGillivray. An evidence-based practice approach to improving nursing care of acute stroke in an Australian emergency department. Journal of Clinical Nursing, 19(1-2):138–144, 2010.

[107] Laurie Lambert, Kevin Brown, Eli Segal, James Brophy, Josep Rodes-Cabau, and Peter Bogaty. Association between timeliness of reperfusion therapy and clinical outcomes in st-elevation myocardial infarction. Journal of American Medical Association, 303(21):2148–2155, 2010.

[108] Kristin N Ray, Amalavoyal V Chari, John Engberg, Marnie Bertolet, and Ateev Mehrotra. Opportunity costs of ambulatory medical care in the United States. American Journal of Managed Care, 21(8):567–574, 2015.

[109] Yu-Kai Lin, Mingfeng Lin, and Hsinchun Chen. Do electronic health records affect quality of care? evidence from the hitech act. Information Systems Research, 30(1):306–318, 2019.

[110] Centers for Disease Control and Prevention. Public health and promoting interoperability programs. Available from `https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms`. Accessed 3/10/2020.

[111] National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey (NHAMCS) Downloadable Documentation, 2017. `ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS`. Accessed 9/12/2020.

[112] Claire Morley, Maria Unwin, Gregory M Peterson, Jim Stankovich, and Leigh Kinsman. Emergency department crowding: a systematic review of causes, consequences and solutions. PloS one, 13(8):e0203316, 2018.

[113] John R Marshall, Robert Katzer, Shahram Lotfipour, Bharath Chakravarthy, Siri Shastry, Jessica Andrusaitis, Craig L Anderson, and Erik D Barton. Use of physician-in-triage model in the management of abdominal pain in an emergency department observation unit. Western Journal of Emergency Medicine, 18(2):181, 2017.

[114] Yosef Berlyand, Ali S Raja, Stephen C Dorner, Anand M Prabhakar, Jonathan D Sonis, Ravi V Gottumukkala, Marc David Succi, and Brian J Yun. How artificial intelligence could transform emergency department operations. The American Journal of Emergency Medicine, 36(8):1515–1517, 2018.

[115] Ziad Obermeyer and Ezekiel J Emanuel. Predicting the future—big data, machine learning, and clinical medicine. The New England Journal of Medicine, 375(13):1216, 2016.

[116] American College of Emergency Physicians et al. American college of emergency physicians. boarding of admitted and intensive care patients in the emergency department. Annals of Emergency Medicine, 38(4):484–485, 2001.

[117] Zlata K Vlodaver, Jeffrey P Anderson, Brittney E Brown, and Michael D Zwank. Emergency medicine physicians' ability to predict hospital admission at the time of triage. The American Journal of Emergency Medicine, 37(3):478–481, 2019.

[118] Susan W Kim, Jordan Y Li, Paul Hakendorf, David JO Teubner, David I Ben-Tovim, and Campbell H Thompson. Predicting admission of patients by their presentation to the emergency department. Emergency Medicine Australasia, 26(4):361–367, 2014.

[119] Bahman Tahayori, Noushin Chini-Foroush, and Hamed Akhlaghi. Advanced natural language processing technique to predict patient disposition based on emergency triage notes. Emergency Medicine Australasia, 2020.

[120] Nicholas W Sterling, Rachel E Patzer, Mengyu Di, and Justin D Schrager. Prediction of emergency department patient disposition based on natural language processing of triage notes. International Journal of Medical Informatics, 129:184–188, 2019.

[121] Filipe R Lucini, Flavio S Fogliatto, Giovani JC da Silveira, Jeruza L Neyeloff, Michel J Anzanello, Ricardo S Kuchenbecker, and Beatriz D Schaan. Text mining approach to predict hospital admissions using early medical records from the emergency department. International Journal of Medical Informatics, 100:1–8, 2017.

[122] Chien-Hua Chen, Jer-Guang Hsieh, Shu-Ling Cheng, Yih-Lon Lin, Po-Hsiang Lin, and Jyh-Horng Jeng. Emergency department disposition prediction using a deep neural network with integrated clinical narratives and structured data. International Journal of Medical Informatics, 139:104146, 2020.

[123] Xingyu Zhang, Joyce Kim, Rachel E Patzer, Stephen R Pitts, Aaron Patzer, and Justin D Schrager. Prediction of emergency department hospital admission based on natural language processing and neural networks. Methods of Information in Medicine, 56(5):377–89, 2017.

[124] Marta Fernandes, Rúben Mendes, Susana M Vieira, Francisca Leite, Carlos Palos, Alistair Johnson, Stan Finkelstein, Steven Horng, and Leo Anthony Celi. Predicting intensive care unit admission among patients presenting to the emergency department using machine learning and natural language processing. PloS one, 15(3):e0229331, 2020.

[125] Bruno P Roquette, Hitoshi Nagano, Ernesto C Marujo, and Alexandre C Maiorano. Prediction of admission in pediatric emergency department with deep neural networks and triage textual data. Neural Networks, 126:170–177, 2020.

[126] Steven Horng, Nathaniel R Greenbaum, Larry A Nathanson, James C McClay, Foster R Goss, and Jeffrey A Nielson. Consensus development of a modern ontology of emergency department presenting problems—the hierarchical presenting problem ontology (happy). Applied Clinical Informatics, 10(3):409, 2019.

[127] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 1135–1144, 2016.

[128] David Alvarez-Melis and Tommi S Jaakkola. On the robustness of interpretability methods. arXiv preprint arXiv:1806.08049, 2018.