

A new method for identifying the primary care treatability of
emergency department visits in a Medicare population

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Molly Moore Jeffery

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Robert L. Kane, Adviser

September 2015

ACKNOWLEDGEMENTS

I owe tremendous thanks to my committee, Bryan Dowd, Jean Abraham, Julian Wolfson, and Bob Kane. You were generous with your time and expertise and are all treasured mentors.

Fernanda Bellolio is a coauthor of the first paper and made the work much better. Frank Wharam provided assistance with the ED indicator procedure list and excellent advice.

I have had wonderful professors and advisors and been invited to work on challenging and meaningful research projects. Thank you to Pinar Karaca-Mandic for exceptionally good academic advising and for her support and mentorship. Thank you to Mary Butler, a brilliant scholar, friend, and boss. I have had a series of outstanding mentors in my academic life, especially Susan Roosenraad, Dolores Peters, Jim Shepard, and Bob Kane; the critical voice I bring to my own work is an echo of their wise and encouraging voices.

My economics study group Ali Bittinger Joglekar, Nikhil Joglekar, Brooke Krause, Aine McCarthy, Travis Smith, and Helen Markelova Trenz, and my office mates Andrea Wysocki, Shri Parashuram, and Ellen McCreedy have been great friends and teachers: I couldn't have done it without you. Thank you to my friends and colleagues at ITFF for encouragement, advice, and community.

Thank you to Julianne Moore, Jim Moore, and Sarah Moore for their support and help with Jack, which made it possible for me to finish a PhD while raising a toddler. I owe more than gratitude to Craig Jeffery for love, patience, and emergency SQL consults. Thank you.

And finally, thank you (again) to Bob Kane, who deserves three separate mentions in the acknowledgments. When I needed employment, data, encouragement (both forceful and gentle, as suited the occasion), editing, letters of recommendation, inspiration, advice, and more RAM, you provided it, with humor and compassion. I am unbelievably lucky to work with you.

DEDICATION

To Craig and Jack, with love

Contents

List of Tables	v
List of Figures	viii
1 Introduction	1
2 Paper One: Development and validation of a new algorithm to identify primary care treatable ED visits in a Medicare population	30
3 Paper Two: Increased primary care visits associated with decrease in primary care treatable ED visits in Medicare population	65
4 Paper Three: Healthcare service use among elderly seasonal migrants	98
5 Conclusions	115
Bibliography	120
A ED indicator procedures	131
B Original Billings algorithm compared to Ballard	143
C Primary care E&M procedures	150

CONTENTS	iv
D HRR groups	151
E Spatial analysis	164
F Paper Three full model results	166

List of Tables

1.1	Summary of Billings/NYU algorithm versions	15
1.2	Mean probabilities of the Billings et al. ED Visit Types by CPT Codes (Table 4 from Wolinsky et al. 2008)	22
2.1	Key Emergency Department Evaluation and Management codes . . .	36
2.2	Summary statistics: validation sample	39
2.3	Key Emergency Department revenue codes	41
2.4	Evaluation and Management algorithm: categorization of Medicare sample ED visits	46
2.5	Comparison of categorization of Medicare sample ED visits: E&M algorithm and Ballard algorithm	49
2.6	Predictive validity of modified E&M Algorithm and Ballard algorithm: hospitalization within 1 day and 1 week from ED visit	52
2.7	Predictive validity of E&M Algorithm and Ballard algorithm: death at 1 week and 1 month from ED visit	54
2.8	Predictive validity of modified E&M Algorithm and Ballard algorithm when the algorithms disagree: death within 1 week and 1 month of ED visit	55

2.9	Predictive validity of standard and modified E&M Algorithm and Ballard algorithm excluding cardiac arrest: death within 1 week and 1 month of ED visit	58
3.1	Summary statistics (2011): HRR groups	84
3.2	Summary statistics: Individual beneficiaries	86
3.3	HRR-level analysis results	88
3.4	Person/year-level analysis: marginal effects (reduction/increase in mean count of outcome per 100 person-years)	93
3.5	Bounded treatment effect analysis results	94
4.1	Sample characteristics	106
4.2	Unadjusted healthcare service use rates (not adjusted for beneficiary characteristics)	108
4.3	Use of Healthcare Services by Migration Status	110
4.4	Predictive margins: incidence rates of ED visits per month of healthcare received	112
A.1	ED Indicator procedure list	131
B.1	Billings/NYU algorithm: categorization of Medicare sample ED visits	144
B.2	Comparison of categorization of Medicare sample ED visits: Billings/NYU algorithm and E&M algorithm	145
B.3	Comparison of categorization of Medicare sample ED visits: condensed Billings/NYU algorithm and Ballard algorithm	147
B.4	Billings and Ballard versions of the algorithm compared	148
D.1	Hospital Referral Region (HRR) groups	151
E.1	HRR-level spatial regression results	165

F.1 Results 167

List of Figures

1.1	Billings/NYU ED classification system	4
1.2	Full Billings/NYU classification scheme	12
2.1	New algorithm	38
3.1	Minnesota algorithm	68
3.2	Dartmouth Atlas hospital referral regions: grouped for use with county-level data.	82

Chapter 1

Introduction

Older people are the most frequent users of Emergency Departments (ED) (Roberts et al. 2008), but research has identified important gaps between elders' needs and the environment of the ED. A 2014 review of the literature on aspects of quality of ED care important to older people found several areas where care sometimes falls short (Shankar et al. 2014). Shankar and colleagues note that "Elderly patients are generally frailer, thus requiring assistance with basic functions such as toileting, walking, and nutrition. Consequently, they find many aspects of the standard emergency care environment uncomfortable."

In addition to being uncomfortable, EDs are also not designed to manage chronic care or non-acute problems. In particular, studies have found that communication and instruction in EDs is poor. Older people (along with general ED patient populations) frequently do not understand the diagnosis, prognosis, and follow-up care plans after an ED visit (in elderly population: Hastings et al. 2012; Hvidt et al. 2014; in general population: Musso et al. 2015)

It is particularly concerning, then, that Kaskie et al. (2011) found that one-third of ED visits by elderly people are categorized as "not severe," indicating that many visits may be treatable in other contexts. These visits could represent significant deficits in provision of care that meets patients' mental, physical, and cognitive needs.

Measuring the primary care treatability of ED visits among Medicare beneficiaries is an important step in developing alternative treatment approaches.

One frequently used technique for classifying ED visits is to apply an algorithm developed by John Billings and colleagues at New York University to categorize visits using the primary diagnosis code recorded in the medical or claims record of each visit. However, the Billings/NYU algorithm may not be well-suited for application in all contexts where it is currently used. In particular, as is further described below, it may be less accurate when applied to populations very different from the one used to create the measure: an all-ages population seen in emergency department in the Bronx, New York.

This study proposes a new measure of primary care treatability of ED visits based on other information and tests whether the new measure performs better than the widely used Billings algorithm in predicting older users' mortality and hospitalizations after the visit. Two applications of the measure look at whether primary care for Medicare beneficiaries reduces primary care treatable ED visits, as measured by the new algorithm, and at the relationship between having local primary care and having primary care treatable ED visits in elderly seasonal migrators ("snowbirds").

This chapter provides background on the Billings/NYU measure which is currently the standard measure used by researchers and policy makers, a review of the literature on measures of ED visit severity and urgency, and a brief discussion of the contribution of the new measure. A full description of the new measure can be found in Chapter Two and is therefore not provided here.

Billings/NYU algorithm

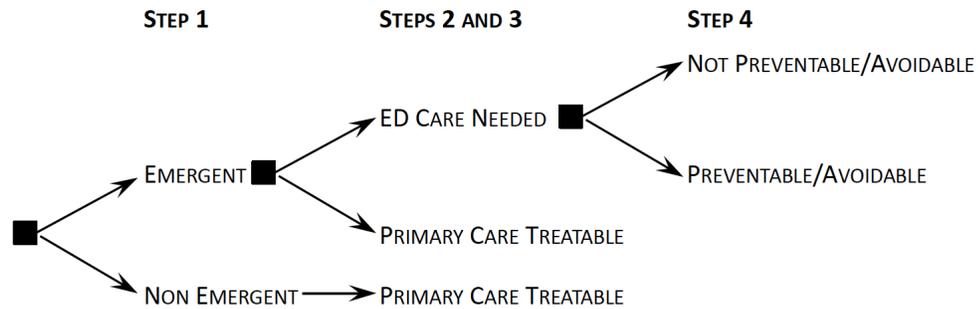
John Billings and his colleagues at New York University proposed using a measure of the primary care treatability of ED visits as an indicator of the performance of

the primary care system and safety net (Gordon et al. 2001). The Billings/NYU algorithm was developed in the 1990s using a sample of 5,700 medical records of ED visits by people of all ages to six hospitals in the Bronx, New York. Physicians reviewed abstracted information from these records and were asked to classify each visit into one of four categories:

1. Nonemergent, indicating medical care was not needed within 12 hours;
2. Emergent/primary care treatable, indicating treatment was required within 12 hours, but could have been performed in a primary care office;
3. Emergent/ED care required but preventable or avoidable, indicating ED care was needed, but the visit could have been prevented with “timely and effective primary care” (for example, an acute episode of asthma); or
4. Emergent/ED care needed, not preventable or avoidable (Billings et al. 2000b).

A visit was considered preventable or avoidable if the primary diagnosis code for the visit is part of the list of ambulatory care sensitive conditions from the AHRQ Prevention Quality Indicators. The Billings approach thus combines multiple elements of the ED encounter: urgency, primary care treatability, and preventability (see figure 1.1).

The categories assigned by the physicians reviewing the medical record abstracts were then tabulated by the primary diagnosis code for each visit to create a percentage of visits with that diagnosis code that fell into each of the four categories (non-emergent–primary care treatable, emergent–primary care treatable, ED care needed–preventable/avoidable, and ED care needed–not preventable/avoidable). Analysts who apply the NYU algorithm to claims data use the primary diagnosis code from their data to match to the NYU diagnosis codes to determine a “likelihood” that the visit falls into each of the four NYU categories.



Source: adapted from Billings et al. 2000c

Figure 1.1: Billings/NYU ED classification system

Though frequently used, the Billings/NYU algorithm may not be appropriate for all situations in which it is currently applied. The algorithm has been used on patient populations very different from the one in the original study. In particular, applying the algorithm to a Medicare population raises significant face validity issues. The population that was used to create the probabilities used in the NYU algorithm included all ages of people presenting to EDs in the Bronx, New York. That reference population would be likely to differ from a national Medicare population in several ways based on co-morbidity, frailty, and geographic variations in practice. The NYU algorithm does not distinguish among age groups: the probability that a primary diagnosis code represents an emergent visit was calculated based on all patients presenting with that primary diagnosis code. This makes the algorithm strongly dependent on the age mix of the original reference population. The difference in the age mix of the Medicare population and that of the NYU reference population raises concerns about applying the NYU algorithm to Medicare data. For a given primary diagnosis code, we would not expect the same level of illness severity if seen in a child, a middle-aged adult, or an older adult. For example, a urinary tract infection (UTI) could have a very different treatment protocol for an 85-year-old woman versus a 25-year-old woman. Similarly, chest pain or confusion have a different clinical meaning for a child

as opposed to an older man.

Another difficulty when using the NYU algorithm is the limited number of diagnoses that it categorizes. The algorithm is limited to the approximately 600 diagnoses that were observed in the analyzed visits in sufficient volume to draw conclusions. Visits with primary diagnoses other than those 600 are left unclassified. If the algorithm is applied to a population experiencing a different set of diagnoses, the results could be biased by excluding those visits. In this sample of ED visits, two of the top twenty diagnoses by volume of visits are not assessed by the NYU algorithm (780.97 altered mental status and 780.60 Fever, unspecified).

In this research, we take a different approach to categorizing ED visits by using procedural rather than diagnostic codes. Most Medicare ED visits (91% of outpatient-only ED visits in this sample) include a physician claim for evaluation and management (E&M), generally assigned by a medical coder based on physician notes in the medical record. There are five procedure codes specific to ED E&M visits, ranging from a visit for a non-emergent, self-limited problem to a high-severity visit for a problem that would cause death or severe morbidity if not attended to quickly. In addition, two critical care E&M codes represent care in the ED for high-severity problems. These seven codes form the basis for the categorization. Intermediate severity visits are examined further for procedures like advanced imaging studies that are likely available only in the ED, not in primary care physician offices.

Variants of this approach have been used before, for example, Davis et al. (2010) categorized visits using the 5 ED-specific E&M codes and the presence or absence of any other billed procedure, and Wolinsky et al. (2008) used just the 5 ED-specific E&M codes. However, the algorithms used in those studies have not been formally validated.

Key terms and constructs

Studies using the Billings algorithm have not been consistent in the language used to describe the construct being measured. The algorithm itself uses three dichotomous terms to describe visits:

1. Emergent versus non-emergent
2. ED care needed versus Primary care treatable
3. Preventable/avoidable versus not preventable/avoidable

The study that formally validated the measure (Ballard et al. 2010) frames the overall construct measured by the algorithm as “severity,” which they describe as “emergent” versus “non-emergent.” But in their version of the algorithm, they ignore the Billings formulation of emergent: the Ballard measure groups together two Billings categories—non-emergent and emergent but primary care treatable—into one category described in the Ballard study as non-emergent. Ballard et al. further assert that the non-emergent claims represent inappropriate ED use, while the emergent claims represent appropriate ED use.

Another study (Lowe and Fu 2008) characterizes the Billings/NYU algorithm as a measure of visit urgency, while attempting to validate it as a measure of changes in access to care.

The proliferation of similar terms being used differently makes it important to be aware of what might be meant by each of these terms.

Emergent/non-emergent, Urgency, Severity: measuring time (how quickly treatment is needed)

The Billings/NYU algorithm defines emergent as requiring medical care within twelve hours (Billings et al. 2000c). Other ED visit measures use different thresholds. The

“immediacy” measure from the ED patient record form for the National Hospital Ambulatory Medical Care Survey (NHAMCS) uses descriptors of visit triage level running from immediate care needed, to emergent, urgent, semi-urgent, and non-urgent. The scale considers a visit emergent if the patient should be seen within 15 minutes; a non-urgent visit is one requiring treatment within 2 to 24 hours (NCHS 2009, 2008). Davis et al. (2010) don’t cite an explicit time period in their definition of urgency. They describe a non-urgent visit as one “that might have been managed outside an ED setting such as in a physician’s office”; low urgency visits could require an ED setting, but are for a “less severe condition” than a high urgency visit. (Note that Davis et al. are apparently using a different definition of “severity” than Ballard et al.) Stedman’s Medical Dictionary defines an emergency as requiring “immediate” treatment, but defines emergent more broadly as “requiring quick judgment and prompt action” (2000); both of these definitions hinge on time as the main criterion.

The official CPT code definitions for the ED E&M codes seems to use a slightly different definition of severity than the one described here that measures time. Table 2.1 below presents the language used in those descriptions to define the five procedures. Severity is reported in a scale ranging from “self-limited or minor” to “high, posing an immediate significant threat to life of physiologic function.” This language describes the seriousness of the problem—the amount of harm it threatens—rather than simply how quickly it must be dealt with. This seems to be what the Davis et al. measure means by severity, as well.

ED care needed/Primary care treatable: measuring resource use

The Billings/NYU algorithm separately considers how quickly treatment is needed and what kind of resources will be needed for treatment. Some visits are considered emergent, but still primary care treatable if they don’t use resources generally unavailable in a primary care setting (for example, an MRI or CT scan).

Davis et al. (2010) consider resource use in their measure of “non-urgent” ED visits. Use of any resources beyond care from a physician or nurse is considered to increase the urgency of an otherwise low-severity visit (where severity here is measured using levels of E&M procedure codes billed). Their reasoning is that any additional procedures “indicate the attending physician sought information during the ED visit that may not have been readily obtained in a physician office setting.” These procedures could include lab tests and cardiac monitoring, but could also include vaccinations and urine pregnancy tests—which are readily available in primary care.

One of the most widely used ED triage instruments—the Emergency Severity Index (ESI)—incorporates likely resource use as an indicator that a patient should be a higher priority for treatment (Gilboy et al. 2011). Patients requiring no resources beyond physician or nurse care are placed in a lower priority triage category, compared to those requiring one or many additional types of resources.¹ The ESI defines resources more exclusively than Davis et al. Lab tests, imaging (X-ray, CT, MRI, ultrasound, etc.), IV medications, and consultations from specialists are considered additional resources, while wound dressing, oral medications, tetanus shots, and crutches, splints, or slings are not.

Preventable, Avoidable: measuring effectiveness of prior care provided

In the Billings/NYU algorithm, “preventable or avoidable” indicates whether the primary diagnosis is included in a list of ambulatory care sensitive conditions (see, e.g., *AHRQ Quality Indicators 2001* for the list)—that is, whether the visit could have been prevented or avoided if the patient had received appropriate primary care prior to the ED visit. The question is not whether the complaint that led to the visit could have been treated in a primary care setting, but rather whether the complaint

1. The triage instrument considers whether the patient requires one or more additional types of resource; multiple different lab tests count as one additional resource.

itself could have been prevented or avoided with good primary care. An example of this construct is an ED visit for an acute asthma attack; if a person with asthma is receiving good primary care for the condition, he will be less likely to suffer from severe attacks that require an ED visit (he will have a rescue inhaler available, or his daily medications will allow him to avoid severe attacks).

This construct has a fairly narrow definition in the context of the Billings/NYU algorithm. It refers only to the short list of conditions that are consistently agreed to be manageable without hospitalizations. The list doesn't purport to be exhaustive; it will miss many individual visits that could have been prevented or avoided with good primary care. Rates of preventable or avoidable visits should be considered to be most interesting in relative comparisons, not as an estimate of the true absolute rate of such visits.

Appropriate/inappropriate

Many of the studies described here use the term "appropriate" to refer to ED visits that meet their urgency, resource use, or severity requirements (see, e.g., Wolinsky et al. 2008; Ballard et al. 2010). However, as Billings et al. (2000c) point out, the appropriateness of a visit also depends on context: "Use of the emergency department for minor conditions may well be rational and appropriate if a patient has no other source of care. Moreover, assessment of urgency by patients can be problematic, and labeling ED use for primary care treatable conditions as inappropriate may misallocate responsibility to the patients themselves." The population used to create the Billings/NYU algorithm had few primary care providers and problems with access during convenient times. EDs are open nights and weekends and don't require appointments, in sharp contrast to primary care practices which may keep bankers' hours and could be fully booked days in advance. An inflexible work schedule could mean that a person would have to take a full day off work for a doctor's office visit;

this could make the ED the only economically feasible source of care. Furthermore, many people are unaware of what is considered to be a healthcare emergency. A broken arm will not receive a high priority for treatment in a moderately busy ED. Other care settings are more “appropriate,” but the general public may have no idea where to go other than an ED. Lacking that knowledge, the person who goes to the ED may be making the most appropriate choice possible.

Development of Billings/NYU ED algorithm

The NYU algorithm was developed by a research team led by John Billings of NYU Wagner Graduate School of Public Service. It was intended to serve as a “gauge of . . . ED utilization patterns” to be used “to understand how changes in the health care delivery system are affecting low-income, uninsured patients” (Billings et al. 2000c). The team was particularly interested in how problems with access to primary care could lead to increased visits to EDs for primary care treatable problems. Before the development of the Emergency Department Profiling Algorithm (as it was called in the initial publications), “the capacity to monitor ED utilization effectively [had] been limited by a lack of data and by methodological challenges. Analysts [had] been able to track overall trends in ED volume but [had] been unable to gain insight into the characteristics of ED use” (Billings et al. 2000c).

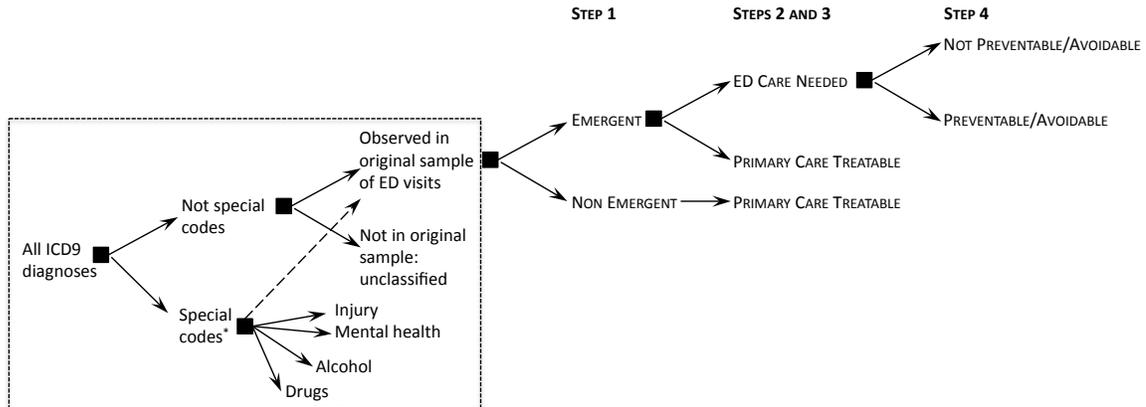
Although the logic behind the design of the new algorithm is not explained in the initial publications associated with it (Billings et al. 2000a, 2000c, 2000b), the basic ideas are clear. Claims data is abundant and relatively inexpensive to procure and use in research, but it doesn’t include as much information about the particulars of a visit as the medical record created during the visit. Medical records are much more detailed, but they are difficult to obtain due to privacy protections and expensive to analyze since much of the information is either unavailable electronically—and must be entered by the researchers by hand—or is difficult to identify because it is buried in

free text fields. The NYU researchers abstracted and analyzed 5,700 medical records of visits to six hospital EDs in the Bronx, New York. They assigned each visit to an algorithm category using the detailed medical record data; then they created a method for applying the information developed in this expensive and time-consuming process to claims data.

Complete ED records were obtained by the research team and the following information was abstracted from each: initial complaint, initial vital signs, age, medical history, procedures performed in the ED, and discharge diagnosis (Billings et al. 2000c). Of the 5,700 total cases abstracted, 3,500 were from 1994 and 2,200 from 1999.

Researchers used the initial complaint, vital signs, medical history, and procedures performed—i.e., information other than the ultimate discharge diagnosis—to determine whether each case “required contact with the medical system within 12 hours,” in which case it was considered emergent; all other cases were classified as non-emergent. Figure 1.2 illustrates the full classification scheme of the final algorithm.

Next, emergent cases were analyzed to determine the “optimal care setting”: primary care or the ED (Billings et al. 2000c). Researchers generally used procedures performed and “resource use” information in this analysis. They note that “no effort was made to assess the appropriateness of the procedures or use of resources” (Billings et al. 2000c). Cases that included only resources and procedures commonly available in primary care were considered primary care treatable. Cases that included the use of procedures generally not available outside the ED were considered to require ED care. In addition, cases with initial complaints that indicate the ED was the appropriate care venue were classified as needing ED care, regardless of procedures performed. The criteria they used for this determination are not provided; it is likely that they were intending to identify visits for complaints like chest pain where it’s



* Special codes that were observed in the original sample of ED visits are also coded for emergence, primary care treatability, and preventability/avoidability in the main part of the algorithm. However, the special codes also include diagnoses *not* seen in the original sample of ED visits.

Figure 1.2: Full Billings/NYU classification scheme

appropriate to go to the ED even though it might turn out to be heartburn, which can, in retrospect, be treated in primary care.

Finally, the discharge diagnosis was used to determine whether cases needing ED care were preventable/avoidable with primary care. The ambulatory care sensitive condition classification (also developed by Billings and researchers at NYU, see Billings et al. (1993)) was used to make this determination. It identifies conditions—like asthma and epilepsy—that should be responsive to good outpatient care to prevent the need for hospitalizations.

In addition to these four ED visit codes, there are five further categories assigned by the Billings/NYU algorithm for “unclassified” and “special” visits (see boxed portion of Figure 1.2). The algorithm includes a set of four special visit types to allow users to separately analyze ED visits for injuries, mental health issues, alcohol-related issues, and drug-related issues. Additional diagnoses fitting those special categories were added to the algorithm at a later date; these additional diagnoses are not categorized in the main part of the algorithm (i.e., as non-emergent, emergent but pri-

mary care treatable, ED care needed–preventable/avoidable, ED care needed–not preventable/avoidable).

Visits whose primary diagnosis code was not assigned to one of the four ED visit codes or one of the special categories are unclassified by the algorithm. These diagnosis codes weren't observed in the initial study sample of ED visits (or weren't observed in sufficient volume to be categorized) and they were not identified as falling within one of the special categories.

Implementation issues

Application of the Billings/NYU algorithm to a dataset requires some programming skill. Billings has provided sample SAS code to make that task easier (Billings 2012). However, there are some inconsistencies between the implementation of the code as written and the published description of the algorithm. The most important difference is in the implementation of the “special” categories of visits.

The initial Commonwealth Fund publications of the algorithm do not include any mention of the special categories (Billings et al. 2000a, 2000c, 2000b). The background section describing the algorithm on Billings's website notes that “users have expressed an interest in examining separately cases involving a primary diagnosis of injury, mental health problems, alcohol, or substance abuse. Accordingly, we have pulled these conditions out of the standard classification scheme, and tabulate them separately” (Billings 2012). The sample code includes a set of diagnosis codes that define each relevant special category (injury, mental health, alcohol-related, drug-related). Some of these codes were part of the algorithm as originally published (i.e., they were observed in the sample of ED visits used to create the algorithm), while others were not. All diagnosis codes that are part of the special categories are removed from the algorithm calculation by the sample code: the code first checks to see whether the diagnosis code related to a visit falls into a special category, if so, it is

flagged for that special category and the probabilities associated with the main part of the algorithm are set to zero. Removing that line of code keeps the special category visits in the main part of the algorithm while retaining the special code information.

Other versions of the Billings/NYU algorithm

The Billings/NYU ED algorithm can be applied in practice in at least four different ways. They differ mainly on two axes: how many of the visit categories are assigned to each visit and which of the visit diagnoses is used to apply the algorithm. Table 1.1 summarizes these approaches.

Official Billings/NYU algorithm application

The official version of the algorithm uses the primary diagnosis code to assign each classifiable ED visit proportionally to the four visit categories or to assign the visit to a special visit category. For example, a visit with the primary diagnosis code 599.0 (Urinary tract infection not otherwise specified) is considered 46.2% likely to be non-emergent, 29.7% likely to be emergent and treatable in primary care, and 24.2% likely to require ED care but be preventable/avoidable with high quality primary care. Thus, each visit with the primary diagnosis code 599.0 would be considered to represent 0.462 non-emergent visits, 0.297 emergent and primary care treatable visits, and 0.242 ED care needed but preventable/avoidable visits.

Greatest probability using principal diagnosis

Some researchers have preferred to assign a single algorithm category to each ED visit in lieu of using the official Billings method of assigning each visit proportionally across the categories. One way of doing this is to use the principal diagnosis code for the visit and assign the category with the greatest probability. In the UTI example

Table 1.1: Summary of Billings/NYU algorithm versions

		Number of categories assigned	
		One	Multiple
Which diagnosis used	Primary	<ul style="list-style-type: none"> • Greatest probability using principal DX 	<ul style="list-style-type: none"> • Official NYU
	Most severe	<ul style="list-style-type: none"> • Greatest probability using most severe DX • Ballard 	
<p>Categories: for Ballard, there are 4 (minor, major, intermediate, special); for the other three versions, there are 5 (non-emergent, emergent–primary care treatable, ED-care needed–preventable/avoidable, ED-care needed–not preventable/avoidable, special)</p> <p>Number of categories assigned: the official NYU algorithm assigns each visit proportionally across multiple categories (e.g., 50% non-emergent, 50% emergent–primary care treatable); the other three versions of the algorithm assign a single category</p>			

above, this would be the non-emergent category. In case of a tie (for example, two categories each assigned 50% probability), researchers may choose to assign the more severe code in the following order: ED care needed–not preventable/avoidable, then ED care needed–preventable/avoidable, then emergent–primary care treatable.

Greatest probability using most severe diagnosis

Although the original design of the algorithm used the primary diagnosis code to categorize each ED visit, the next two versions of the NYU algorithm consider all diagnosis codes assigned to a visit. In this version of the application of the algorithm, the NYU algorithm is applied to every diagnosis attached to an ED visit claim; the most severe category assigned to any of these diagnoses is then selected as the category for the visit. For example, if a visit were assigned two diagnosis codes including a primary diagnosis of 922.1 (contusion of chest wall), which is 62.3% likely to be emergent but primary care treatable according to the algorithm, and a secondary diagnosis of 786.05 (shortness of breath), which is 60% likely to be ED care needed–not preventable/avoidable, that visit would be classified as ED care needed–not preventable/avoidable. This variant of the algorithm has the benefit of leaving

fewer visits unclassified, since it requires that just one of potentially ten or more diagnoses assigned to the visit be classified by the algorithm.

Ballard et al. version of algorithm

The Billings/NYU algorithm had been widely used but never formally validated until the 2010 Ballard et al. publication. The researchers used a modification of the original algorithm in their validation study. As in the “greatest probability using most severe diagnosis” version, the Ballard version of the algorithm considers all diagnoses assigned to an ED visit and assigns the most severe algorithm result to the visit. However, the Ballard version consolidates the four Billings/NYU categories into three: minor, intermediate, and major.

First, for each diagnosis, the probabilities of the two least severe categories (non-emergent and emergent, but primary care treatable) are summed, and the probabilities of the two most severe categories (ED care needed–preventable/avoidable and ED care needed–not preventable/avoidable) are summed. If the sum of the two least severe categories is greater than 50%, the “minor” category is assigned. If the sum of the two most severe categories is greater than 50%, the “major” category is assigned. If both sums are 50%, the “intermediate” category is assigned. Finally, the most severe result (first major, then intermediate, then minor) across all visit diagnoses is assigned to the visit.

The sample used to validate the algorithm included three years of data from an integrated delivery system (IDS) with approximately 2.8 million members (10% Medicare, 90% commercial members). The researchers excluded members who “regularly received part of their care outside the IDS” (5.5% of members) and infants under one year old (1.2% of members) (Ballard et al. 2010). The outcome measures for which the algorithm was validated were hospitalizations and deaths (both in and out of the delivery system). They observed 1.1 million ED visits with ICD-9 codes (excluding

35,000 visits that had only a mental health or substance abuse diagnosis), of which 94.3% had at least one ICD9 code classified by the NYU algorithm.

The Ballard team’s approach to validating the NYU algorithm was to describe the predictive validity of the algorithm for future hospitalizations (on the same day, within one day, and within seven days of the ED visit) and for death within thirty days of the ED visit (Ballard et al. 2010). They used a generalized estimating equation approach to estimate odds ratios of the outcome measures for patients with nonemergent (reference level), intermediate, and emergent ED visits.

They state that they used the `xtlogit, pa` command in Stata; they were therefore estimating the following model:

$$\text{logit}(E(y_{it})) = \beta \mathbf{x}_{it} + \gamma_1 g_{it,\text{int}} + \gamma_2 g_{it,\text{emerg}}, y \sim \text{Bernoulli} \quad (1.1)$$

$$\text{with working correlation matrix } \mathbf{R}_{t,s} = \begin{cases} 1 & \text{if } t = s; \\ \rho & \text{otherwise} \end{cases}$$

where y_{it} is 1 if subject i experienced the outcome y associated with ED visit it and 0 otherwise; \mathbf{x}_{it} is a vector of person- and visit-specific variables including age, gender, neighborhood socio-economic status derived from the 2000 Census, medical care utilization from the previous year (ED visits, office visits, and hospitalizations), comorbidity information using the diagnostic cost group score, ED and drug cost-sharing levels, whether subject i had a regular primary care physician, indicator variables for month and year of the ED visit, and a fixed effect for the medical center of the ED. The independent variables of interest are indicators for whether ED visit it was categorized by the algorithm as an intermediate visit ($g_{it,\text{int}}$) or an emergent visit ($g_{it,\text{emerg}}$); the omitted category is a non-emergent visit. Separate models were run for commercial and Medicare patients.

The analysis showed that both intermediate and emergent ED visits were predictive of all four of the outcomes (hospitalizations within 0, 1, and 7 days and thirty-day mortality) for both the commercial and Medicare populations compared to non-emergent ED visits. For example, the odds of death within 30 days of the ED visit are 1.7 (CI 1.6–1.8) times higher for an intermediate visit versus a non-emergent one and 2.2 (CI 2.1–2.3) times higher for an emergent visit in the Medicare population. The results are similar for the commercial population.

These results serve as an important validation of the algorithm, but the particulars of the analyses may not address the purposes for which most researchers use the Billings/NYU algorithm.

The Medicare sample used in this study were members of Kaiser Permanente–Northern California, a managed care plan. Most studies of Medicare managed care plans have shown favorable risk selection into these private plans (see, e.g., Mello et al. 2003)—that is, people enrolled in Medicare managed care plans tend to be healthier than those enrolled in Fee for Service Medicare (FFS). Indeed, the study sample may be significantly healthier than the general Medicare population. The study reports that just 9.2 percent of people in the Medicare sample had any hospitalizations in 1998. By comparison, the 1998 rate of hospital discharges for Medicare FFS enrollees in California was 31.3 per 100 beneficiaries (Dartmouth Atlas 2015) and the nationwide proportion of Medicare FFS enrollees with at least one hospitalization in 2000 was 23.3 percent (MedPAC 2012).

If the study sample is different from the general Medicare FFS population (the majority of all Medicare beneficiaries) in underlying health or in medical care use, the validation of the Billings/NYU measure in the Kaiser Permanente sample may not generalize to the broader Medicare population. This is important because most studies using Medicare claims data use only the FFS population, since claims data for the managed care population is incomplete.

The Ballard study, then, may not have adequately validated the NYU algorithm for the general Medicare population.

Other approaches to categorizing ED visits

Diagnosis code approaches

A measure of diagnoses representing potentially avoidable ED visits was created for a quality improvement project undertaken by the Medi-Cal division of the California Department of Health Care Services (DHCS) to reduce the number of avoidable ED visits among Medi-Cal managed care beneficiaries (Delmarva Foundation 2008). Little detail is available on the creation of this measure beyond the brief description in the 2008 report:

This HEDIS-like measure developed by DHCS and the participating plans summarizes the percentage of designated “avoidable” ER visits. To develop [this measure], the collaborative discussed methods to identify and measure avoidable visits during weekly meetings for eight months in 2007. The collaborative reviewed published literature and consulted with noted experts on ER use from the University of California at San Francisco, the University of California at Davis, and New York University for assistance in developing a practical list of diagnosis codes for selected avoidable visits.

The Washington State Hospital Association (WSHA) subsequently used the Medi-Cal measure in its own report on potentially avoidable ED use (WSHA 2011) and included the full list of diagnoses in the measure.

Three categories of diagnoses are included: a list of 165 diagnoses considered to represent avoidable ED visits, a list of 69 dental diagnoses, and a list of 356 mental health and substance abuse diagnoses. Only diagnoses on the first list are counted as

avoidable ED visits. The other two lists are for identifying other ED visits of interest to the group that created the measure.

The group developing the measures seems to have focused on specificity over sensitivity in selecting the diagnoses for the list of avoidable ED visit indicators. The 165 diagnoses represent a relatively small number of types of problems: general medical exams (follow-ups, pre-op exams, etc.), conjunctivitis, bladder infections, chronic and acute ear/nose/throat issues including otitis media and minor upper respiratory tract infections, minor backaches and headaches, and itching/rashes. The vast majority of people coming to the ED for these problems will certainly have been treatable in less acute care settings. However, this list will also leave out large numbers of potentially primary-care treatable ED visits where the presenting problem could represent a range of severities—like stomach pain, injuries, etc.

For diagnoses that are included in both the Medi-Cal list and the Billings/NYU algorithm, the results of the two algorithms are similar. Of the 165 diagnoses on the Medi-Cal list, 62 are fully categorized by the Billings/NYU algorithm; most (80%) of the total probability assigned to those 62 diagnoses is non-emergent, 16% is emergent but primary care treatable, 2% is ED care needed—preventable/avoidable, and 2% is ED care needed—not preventable/avoidable. Of the remaining 103 diagnoses on the Medi-Cal list, one is rated only as a special category (V704 Medico-legal exam, which Billings/NYU categorizes as alcohol related). The rest are unclassified by the Billings/NYU algorithm (though many of them are closely related to diagnoses that are classified by the NYU algorithm).

Procedure code approaches

The set of five ED-specific E&M codes have been used before to classify ED visits. Wolinsky et al. (2008) grouped the two lower intensity codes (99281–2) and the three higher intensity codes (99283–5) to study ED use patterns among Survey on Assets

and Health Dynamics Among the Oldest Old (AHEAD) respondents. They used four approaches to informally validate their measure of ED visit intensity. First, they used exploratory factor analysis to verify that the grouping of the codes was supported empirically. They found that the annual volume of ED visits falling under each of the five codes did, in fact, load on the two hypothesized factors (in plain language, the two lower intensity codes trended differently over time than the three higher intensity codes). Next, they looked at some potential indicators of the criterion validity of the CPT codes being used to measure “ED appropriateness.” Their reference variables were

1. Percentage of all ED visits
2. Percentage arriving by ambulance
3. Percentage admitted to hospital
4. Mean length of stay if admitted
5. Mean medical charges if admitted
6. Percentage dying within 2 years of baseline AHEAD interview, using all-cause NDI mortality

Items 2, 3, and 5 showed the expected monotonically increasing pattern across the five CPT codes (i.e., 99285 had the highest rate or mean, 99284 the next highest, and so on down to 99281).

The third approach to verifying their ED visit measure was to classify the visits using the Billings/NYU algorithm. Their results are worth replicating here (see Table 1.2).

The authors note that, with two exceptions, the data show a monotonic relationship in the expected direction between visit intensity as measured by E&M CPT

Table 1.2: Mean probabilities of the Billings et al. ED Visit Types by CPT Codes (Table 4 from Wolinsky et al. 2008)

Billings category	99281	99282	99283	99284	99285
Non-emergent	0.214	0.195	0.159	0.118	0.080
Emergent visit, primary care treatable	0.294	0.330	0.298	0.247	0.207
ED care needed, but the condition was preventable/avoidable	0.062	0.090	0.103	0.153	0.189
ED care needed, and the condition was not preventable/avoidable	0.250	0.199	0.253	0.290	0.353
Not classified	0.180	0.186	0.187	0.192	0.171
Total	1.000	1.000	1.000	1.000	1.000

codes and visit “appropriateness” (their term) as measured by the Billings algorithm ($\chi^2 = 213.11; p < 0.0001$) and that the two discrepancies both involve the lowest volume category: 99281. Specifically, the emergent visit, primary care treatable proportion for 99281 visits is lower than for 99282 (more intense) visits and the ED care needed, not preventable/avoidable proportion for 99281 visits is higher than for 99282 visits. There is not enough information provided in the article to determine which proportions are statistically distinguishable from each other; the sample size of ED visits is not given, but there are 4,135 respondents included in the study, only 44 percent of whom had any ED visits in the 4-year period.

Interestingly, the authors do not mention the high level of disagreement they found between Billings categories and visit intensity. Fully one-quarter of the least intense (99281) visits were rated by the Billings algorithm as ED care needed, not preventable/avoidable, while almost 30 percent of the highest intensity ED visits (99285) were rated as primary care treatable by the Billings algorithm (non-emergent + emergent, primary care treatable).

The final approach to verifying the Wolinsky et al. ED visit measure was to compare the five most frequent diagnoses for low intensity (99281–2) visits versus high intensity (99283–5) visits. They found that the high intensity visit diagnoses

were very different from the low intensity visit diagnoses and seemed to represent much more severe conditions.

The Davis et al. 2010 study uses a very similar algorithm to the Wolinsky study. The Davis team also groups the two lower “severity”² visit codes (99281–2) and the three higher severity visit codes; however, they add a third category of visits, splitting the low severity visits by whether any further procedures were billed beyond the E&M code. So for the Davis study, low severity visits with no other procedures billed were considered “low urgency,” low severity visits with at least one additional procedure billed were considered “medium urgency,” and high severity visits (CPT codes 99283–5) were considered “high urgency.”

The Davis study does not include an attempt to validate their visit urgency measure. They don’t cite the Wolinsky et al. study, and may not have been aware of it. They note that their definition of low urgency visits was selected using “both expert opinion and evidence from administrative data that no care was required beyond that available in a physician office.” The presence of additional billed procedures beyond the E&M code is meant to indicate “the attending [ED] physician sought information during the ED visit that may not have been readily obtained in a physician office setting.” However, they note that the additional billed procedures included vaccinations, which presumably would be available in a primary care setting; they may have improved their measure somewhat had they attempted to distinguish between additional procedures that are clearly available in primary care and those that may not be.

A 2010 study by Kaskie et al. (with Wolinsky as the senior author, along with several of the other Wolinsky et al. (2008) authors) brings together the diagnosis-based and procedure code-based approaches to categorizing ED visits. The authors

2. As noted above, there is no agreement in the literature over what the CPT codes measure—severity, intensity, urgency, etc. In discussing each study’s approach, I’ve used the same term as the authors.

argue that the two approaches measure different dimensions of ED use. They used an approach similar to the Ballard et al. 2010 version of the Billings/NYU algorithm to measure clinical severity of an ED visit; as with the Ballard study, NYU algorithm probabilities were summed for each visit. In this case, instead of using a 50% cut-off for distinguishing between more and less severe visits, the Kaskie team used a 75% cut-off. In other words, if the sum of the probability of ED care needed, preventable/avoidable and the probability of ED care needed, not preventable/avoidable was greater than or equal to 75%, the visit was classified as a severe visit. If that sum was less than or equal to 25%, the visit was classified as non-severe. If the sum was greater than 25% and less than 75%, the visit was considered indeterminate. The study authors note that they took this approach from Wharam et al. (2007); the Ballard team (also publishing in 2010) used the 75% cut-off and a 90% cut-off as sensitivity checks in their validation study.

The Kaskie study used the ED-specific E&M codes to measure the intensity of the care provided in the visit. They grouped the codes into two categories: highest intensity visits with E&M code 99285 versus all other visits (99281–4). Some analyses were also conducted using all five E&M codes separately.

The two measures—one based on diagnosis codes and the other on procedure codes—were tested for concordance using 15 years of data on the AHEAD 1993 cohort (1991–2005). They found the two measures showed very low agreement; when the two binary measures (highest severity vs. others for the NYU algorithm and highest intensity vs. others for CPT codes) were cross-classified, the kappa statistic was 0.18 (CI 0.167–0.196), indicating “slight” agreement, using the Landis and Koch (1977) guidelines.

They compared the predictive validity of the two binary measures on hospitalization “following the ED episode”; it isn’t clear from the text what time frame is used, but it seems to be within 3 days of the ED visit. They found that the CPT-based mea-

sure had higher predictive power than the modified NYU algorithm measure based on diagnoses both on simple odds ratio calculations (using adjusted odds ratios to identify the independent effects of the two measures) and using logistic regression to adjust for covariates including age, race, smoking status, geography, and time period. In the regression model, the highest severity modified-NYU code had 3.31 times higher odds of hospitalization compared to all other severity visits, while the highest intensity CPT code had 5.70 times higher odds of hospitalization compared to all other intensity visits. The interaction of the two—visits with highest intensity and highest severity—had 8.09 times higher odds of hospitalization compared to all other visits.

The authors assert that the two measures are picking up “two distinct constructs” (presumably severity and intensity) that may be “differentially useful in predicting other outcomes such as continuity of care, nursing home placement, or death.” They also note that they could vary in their predictive validity in different populations.

The contribution of this measure

Many stakeholders in health care would like to define exactly which ED visits represent appropriate use of health care resources. What exactly is meant by appropriate use is difficult to pinpoint. The central question is whether the problem could have been handled in a less resource-intensive setting, like primary care. But the problem is more complicated than it might seem.

- How should access restrictions be factored into the determination of the most appropriate venue for care? If a patient’s primary care physician is booked out 3 weeks, is the ED the right place for an urgent but not severe problem like a UTI? What if the patient works during regular office hours or has access to transportation only in the evenings and on weekends?

- What about symptoms that could have indicated a severe problem but are later found to be benign? For example, chest pain that turns out to be heartburn or rectal bleeding that turns out to be hemorrhoids? Surely the ED is not the best place to treat heartburn and hemorrhoids, but the patient may have no way of knowing the cause of many non-specific and worrying symptoms.

Thus, though we are most interested in determining which visits represent appropriate use of the ED, we see that visit appropriateness must be judged by what the patient knows at the time he chooses where to present and must take into account numerous contextual attributes like time of day, day of the week, and presence and accessibility of alternate venues of care.

Unfortunately, few of these contextual elements are observable in administrative data. In general, of all these elements, we can see only the day of the week. Time of day is not recorded in claims data. The reason for the visit is generally not recorded, as the diagnoses present on a Medicare outpatient claim should represent the final diagnosis at the end of the visit (CMS 2012a, Ch. 23.10A). The presence of alternate venues of care like urgent care, retail clinics, and extended primary care office hours could in theory be collated from public sources, but it would be a massive undertaking. And the accessibility of alternate venues: primary care booking density, financial incentives from supplemental insurance coverage, and availability of transportation are nearly impossible to ascertain without surveying patients directly.

What this measure addresses

Of all these key concepts, the new algorithm measures only one: primary care treatability, that is, the resources used in the visit. A second concept, severity, is partially addressed by the measure. Resources used in an ED visit can be measured or inferred from FFS claims with a reasonable amount of accuracy, particularly when they are

billed for directly, as with imaging, lab tests, medications given, and so on. Resources that cannot be directly billed for such as physician and nurse time can be inferred through the E&M code billed for a visit. The rules governing which level of E&M code can be billed include criteria relating to the complexity of decision making by the physician and to the thoroughness of the examination and history-taking, which can be considered rough proxies for the amount of time spent with the patient.

The severity of a visit is partially addressed by the new algorithm, but only to the extent that the procedures performed indicate the physician's belief about the severity or potential severity of the visit. For example, the E&M code requirements for reimbursement include a characteristic severity of the presenting problem. The code 99281 is stated to be "usually" for a "self-limited or minor" presenting problem, while the code 99285 "usually" suggests "an immediate significant threat to life or physiological function."

The ED indicator procedure list provides another source of information about the severity of the visit. If cardiac enzyme labs are ordered, we could infer that the physician believed there to be a risk of heart attack, suggesting a potentially severe visit, regardless of whether the final diagnosis was a heart attack. The procedures on the ED indicator list were selected because they are performed more frequently in the ED than in primary care and they fulfill at least one of the following criteria:

1. Performance of the procedure always or nearly always indicates a visit requiring ED care; examples: Endotracheal intubation, CPR;
2. Procedure is unlikely to be available in a primary care office and performance is likely to be time-sensitive; examples: CTs and MRIs, morphine injections;
3. Procedure is appropriate and widely available in primary care, however performance in the ED indicates the physician believed there may be a severe problem; examples: lab tests for troponin or creatine kinase.

What this measure cannot address

The question of whether procedures present in ED claims data were necessary is an important, but likely currently unanswerable question. Despite the nearly universal belief that some health care delivered is unnecessary, in most cases there is little agreement on exactly which procedures should not have been performed. Recent work to identify overuse has focused on carving out a list of narrowly defined procedures and populations that suggest inappropriate use. For example, the Choosing Wisely campaign distributes lists created by 67 physician and nurse specialty organizations with 5–15 procedures on each list that are not supported by medical evidence; i.e., they probably represent inappropriate use of resources. An example from the American College of Emergency Physicians (ACEP) is “Avoid computed tomography (CT) scans of the head in emergency department patients with minor head injury who are at low risk based on validated decision rules.” (ACEP 2013). It would be difficult to determine from claims data whether a particular CT scan in a case of minor head injury represents overuse under this guideline. The decision rules are not recorded in claims data. Of the ten Choosing Wisely potentially overused procedures identified by ACEP, none could be confirmed from claims data alone. Future research with datasets linking EHRs to claims data may allow us more clarity on overuse of health care services that could inform future versions of the new ED algorithm.

Because physicians and hospitals are paid based on the level of E&M code they bill, we might expect to see significant upcoding: that is, billing for a higher level of E&M code than is indicated by the work actually done. The flip side of this is that Medicare has an incentive to audit procedure coding. In fact, upcoding of E&M services has been the subject of audits of Medicare data. See paper one “Limitations” for a more thorough discussion. The evidence is mixed on how much of a problem upcoding represents in ED E&M codes. However, if there is a significant problem, it seems to be dominated by coding errors of one level (for example, a visit with a “true”

E&M code of 99281 is billed as 99282 or 99284 is billed as 99285). The design of the new algorithm makes it relatively robust to this level of coding error. Furthermore, the validation results are strong for the actual E&M codes billed. If E&M coding patterns change in the future, validation may need to be repeated and the measure adjusted.

Chapter 2

Paper One: Development and validation of a new algorithm to identify primary care treatable ED visits in a Medicare population

The Institute of Medicine has described emergency care as “a window on health care, revealing both what is right and what is wrong with the care delivery system” (IOM 2007). Emergency departments (EDs) provide care to all comers, including those on the brink of death and those with less urgent problems. However, EDs are an inefficient, expensive way to deliver non-emergent care. Greene (2014) reports a Healthcare Bluebook estimate of a fair price for a minor ED visit as \$637. In contrast, the Bluebook fair price for a new patient visit to a physician office for a minor problem is \$82 (Healthcare Bluebook 2014).

EDs are not just expensive, many are also crowded and likely to become more so as the Affordable Care Act is fully implemented (see, e.g., McClelland et al. 2014). ED overcrowding may reduce the quality of care and negatively affect patient outcomes (Schiff 2011; Felton et al. 2011; LaMantia et al. 2010; Olshaker 2009; Ruger et al. 2012).

Though people of all ages use the ED, elderly people are particularly heavy users.

Medicare beneficiaries have the highest visit rate of any age group (Roberts et al. 2008). Their visits tend to be resource intensive (Hampton 2008). The consequences of these visits are not just the potential for overcrowding and higher expenses; ED visits can also be taxing both mentally and physically for an ill or frail elder. Furthermore, Kaskie et al. (2011) found that one-third of their visits are categorized as “not severe,” indicating that many visits may be treatable in other contexts.

For all these reasons, it is important to consider when an ED is the best venue for a medical encounter. An ED visit that could have been treated in primary care may represent a significant waste of resources, for both the patient and the health care system. Identifying these ED visits among Medicare beneficiaries is thus a high priority for researchers and policy makers seeking to understand why these visits happen and how to help patients determine the best place to go when in need of medical care.

One frequently used approach to identify ED visits that could have been treated in primary care relies on an algorithm developed by Billings and his colleagues to categorize visits using the primary diagnosis code recorded in the medical or claims record of each visit (Billings et al. 2000a).

Though frequently used (see Jones et al. (2011) for a partial list of its use in recent research studies and in program development), the Billings/NYU algorithm may not be appropriate for all situations in which it is currently applied, especially Medicare. The algorithm has been used on patient populations very different from the one in the original study. In particular, applying the algorithm to a Medicare population raises significant face validity issues. The population that was used to create the probabilities used in the NYU algorithm included all ages of people presenting to EDs in the Bronx, New York. That reference population likely differs from a national Medicare population in several ways based on co-morbidity, frailty, and geographic variations in practice. The NYU algorithm does not distinguish among age groups:

the probability that a primary diagnosis code represents an emergent visit was calculated based on all patients presenting with that primary diagnosis code. This makes the algorithm strongly dependent on the age mix of the original reference population. The difference in the age mix of the Medicare population and that of the NYU reference population should make us wary of applying the NYU algorithm to Medicare data. For a given primary diagnosis code, we would not expect the same level of illness severity if seen in a child, a middle-aged adult, or an older adult. For example, a urinary tract infection (UTI) could have a very different treatment protocol for an 85-year-old woman versus a 25-year-old woman. Similarly, chest pain has a different clinical meaning for a child as opposed to an older man.

In fact, there are clear indications that the NYU algorithm's underlying assumptions are not true for a Medicare population. Of all diagnoses coded by the NYU algorithm, 68% are determined to be 100% primary-care treatable—that is, all patients in the original NYU sample who had that primary diagnosis were determined to have been treatable in primary care.¹ In our sample, we observed 148,000 ED visits with a primary diagnosis considered 100% primary-care treatable by the NYU algorithm. In 49% of these visits, the patient was admitted to the hospital, received critical care in the ED, or received the highest/most severe level of ED evaluation and management code.²

Another difficulty when using the NYU algorithm is the limited number of diagnoses that it categorizes. The algorithm is limited to the approximately 660 diagnoses that were observed in the analyzed visits in sufficient volume to draw conclusions.³

1. This includes both the non-emergent code and the emergent but primary care treatable code in the NYU algorithm.

2. Looking only at the diagnoses rated 100% non-emergent in the NYU algorithm, we observe 48,000 ED visits of which 28% were admitted to the hospital, received critical care in the ED, or received the highest/most severe level of ED evaluation and management code.

3. There is some grouping of diagnoses in the NYU algorithm code, so it is most accurate to say that there are 659 diagnosis groups coded by the NYU algorithm. Some groups are ICD9 codes grouped at the 3-digit level, others cross multiple 3-digit ICD9 codes. In our sample, 966 distinct 3

Visits with primary diagnoses other than those 660 are left unclassified. If the algorithm is applied to a population experiencing a different set of diagnoses, the results could be biased by excluding those visits. In this study, the NYU algorithm left 24% of all ED visits unclassified.

Because the Billings algorithm is based on diagnoses, it may be best used on relatively heterogeneous populations. For example, it would be difficult to study the severity of ED visits in a group of people with congestive heart failure participating in a care coordination program because the entire group has the same severe diagnosis. Relatedly, the Billings algorithm doesn't allow visits with the same diagnosis to vary in severity. This likely reduces the ability of the algorithm to measure the effect of policy changes, since, in order to see a change in the ED visit severities as measured by the algorithm, you would have to see a change in the diagnosis mix seen in the ED, not just a change in the true severity of visits.

The form of the original NYU algorithm makes it difficult to use in regression analyses. It is a set of four proportions, summing to one, with a substantial number of zeros, making it an especially tricky example of compositional data. The techniques used to analyze compositional data generally require awkward transformations—usually log ratios (Bacon-Shone 2011)—that complicate interpretation of regression results and require imputation of positive values to replace the zeros (Martín-Fernández et al. 2011). As a result, analysis using the original NYU algorithm is generally limited to tabulations of visits, without the ability to use regression to adjust for covariates.

Researchers have developed alternative versions of the algorithm to overcome these problems. The Ballard version of the algorithm, the only version to be formally validated, considers all diagnoses assigned to an ED visit and assigns the most severe algorithm result to the visit (Ballard et al. 2010). The Ballard version also consol-

to 5 digit ICD9 codes are classified by the algorithm, representing over 2 million visits. A further 1,743 diagnosis codes observed in our sample in 130,000 visits are classified only as related to mental health, drugs, alcohol, or injury (with no emergence rating).

idates the four Billings/NYU categories into three: nonemergent, intermediate, and emergent. First, for each diagnosis, the probabilities of the two least severe categories (non-emergent and emergent, but primary care treatable) are summed, and the probabilities of the two most severe categories (ED care needed–preventable/avoidable and ED care needed–not preventable/avoidable) are summed. If the sum of the two least severe categories is greater than 50%, the “nonemergent” category is assigned. If the sum of the two most severe categories is greater than 50%, the “emergent” category is assigned. If both sums are 50%, the “intermediate” category is assigned. Finally, the most severe result (first emergent, then intermediate, then nonemergent) across all visit diagnoses is assigned to the visit. This version solves some of the problems described above, including the number of visits unclassified and the compositional data issue. The problems inherent to the diagnosis approach remain.

This study: 1. Proposes a new measure of primary care treatability of ED visits that solves some of the problems presented in applying the Billings/NYU algorithm; and 2. Uses Medicare data to compares the new measure’s performance to the established Ballard variant of the Billings/NYU algorithm in predicting hospitalization and death following an ED visit.

Methods

Study rationale and approach

In this research, we categorize an element of the ED visit classification, namely, the primary care treatability of Medicare patients’ ED visits, by using procedure codes. There are five procedure codes specific to ED E&M visits, ranging from a visit for a non-emergent, self-limited problem to a high-severity visit for a problem that would cause death or severe morbidity if not attended to quickly. In addition, two critical

care E&M codes represent care in the ED for high-severity problems. See Table 2.1 for descriptions. These seven codes form the basis for the new algorithm’s categorization. Most Medicare ED visits—91% all ED visits in sample of Medicare claims used in this study—include a physician or institutional claim for one of these codes. Intermediate severity visits are examined further for procedures like advanced imaging studies that are generally not available in primary care physician offices to determine whether they were likely primary care treatable. Figure 3.1 presents the approach as a flowchart.

Table 2.1: Key Emergency Department Evaluation and Management codes

E&M Code	Requirements for reimbursement			Characteristics	
	History	Exam	Medical Decision Making	Counseling/Coordination of care	Severity of presenting problem
99281	Problem focused	Problem focused	Straight-forward	Consistent with the nature of the problem(s) and/or family's needs	Usually self limited or minor
99282	Expanded problem focused	Expanded problem focused	Low complexity	Consistent with the nature of the problem(s) and/or family's needs	Usually low to moderate
99283	Expanded problem focused	Expanded problem focused	Moderate complexity	Consistent with the nature of the problem(s) and/or family's needs	Usually moderate
99284	Detailed	Detailed	Moderate complexity	Consistent with the nature of the problem(s) and/or family's needs	Usually high, requiring urgent evaluation by physician, but not posing an immediate significant threat to life or physiological function
99285	Comprehensive	Comprehensive	High complexity	Consistent with the nature of the problem(s) and/or family's needs	Usually high, posing an immediate significant threat to life or physiologic function
Critical care	Description			Requirement for reimbursement	
99291	Evaluation and management of critically ill or critically injured patient; first 30–74 minutes			Constant physician attention; high complexity decision making to assess, manipulate, and support vital system function to prevent or treat single or multiple vital organ system failure. [†]	
99292	Each additional 30 minutes				

Source: All text taken directly or paraphrased from official AMA CPT descriptions except [†] from CMS Pub 100.

In cases where a visit generated multiple E&M codes, the following rules were used to select the code used to classify the visit:

1. E&M codes charged by physicians were selected over those charged by institutions. The rules for institutions choosing which E&M code to charge are less standardized than those for physicians.
2. If there were multiple codes charged by physicians (or by the institution in the absence of a physician E&M code), the most severe code was used.

Variants of this approach have been used before, for example, Davis et al. (2010) categorized visits using the 5 ED-specific E&M codes and the presence or absence of any other billed procedure, and Wolinsky et al. (2008) used just the E&M codes.

Population studied

Two years of claims data (2011-2012) for a representative sample of Medicare beneficiaries was used to develop and validate the new measure. The Limited Data Set files made available by the Center for Medicare & Medicaid Services include claims for nearly all services paid for by Medicare (Part D drug claims are not available) including inpatient, outpatient, physician services, home health, durable medical equipment, hospice, and skilled nursing facilities. Information that could be used to identify individual beneficiaries has been stripped from the files according to HIPAA privacy rules. The data includes limited demographic information for all sampled beneficiaries, regardless of whether they received any medical care: beneficiary age (top-coded at 100 years of age), county of residence, gender and race, the reason for receipt of Medicare benefits (old age, end-stage renal disease, etc.), and indicators for receipt of state subsidies that serve as a proxy for dual eligibility for Medicare and Medicaid; the claim files include unencrypted physician and institution identifiers, procedure codes, and diagnosis codes, as well as submitted and paid charges. A nationwide five

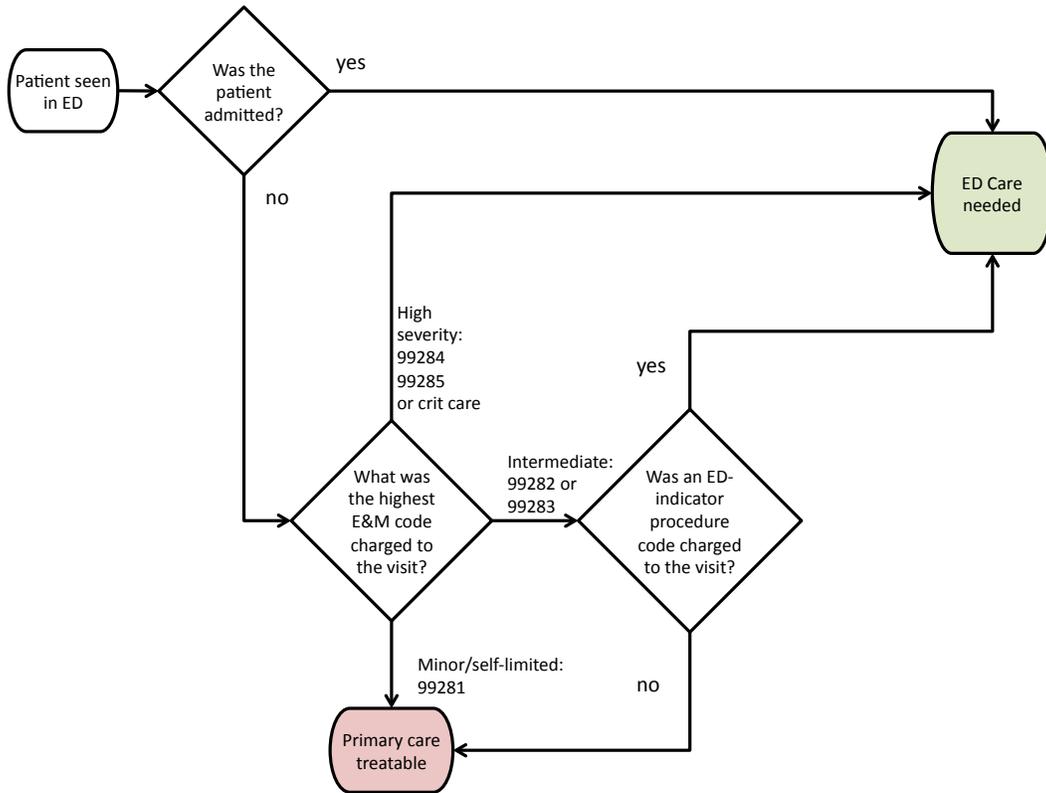


Figure 2.1: New algorithm

percent sample is included in the files; the sample is selected based on the Medicare identification number, ensuring that beneficiaries remain in the sample as long as they are receiving Medicare benefits. See CMS 2014 for further details.

Table 2.2: Summary statistics: validation sample

	Mean	SD	Min	Max
Age (in 2011)*	67.75	14.51	-1	98
Race				
white	82.36%			
black	12.07%			
other	5.571%			
Female	57.32%			
Medicaid**	27.92%			
HCC community score	2.507	1.656	0.121	18.28
Died during study period	18.67%			
ED visits in sample	2.957	4.033	1	351
Months in sample	21.86	5.212	1	24
Total ED visits	2,448,112			
Total beneficiaries	827,844			

* Three beneficiaries born in 2012 have an age of -1 in 2011.

** Operationalized as any months of state coverage buy-in from denominator file.

Identifying ED visits in Medicare claims files

Emergency Department (ED) visits can appear in multiple Medicare claims files. In general, we expect to see at least two claims associated with each visit: an institutional claim and a physician claim. The first is submitted by the hospital in which the ED is located, while the second is submitted by the physician who treated the patient.

Using the Limited Data Set (LDS) for 2011 and 2012, we identified and matched all emergency department visits across the institutional and carrier files by patient ID and date, taking into account Medicare payment rules requiring bundling of outpatient

and inpatient charges within 3 calendar days (the three-day payment window—further detail below). Note that there is no claim-level identifier that definitively groups claims in different files as belonging to the same ED visit—the user must create decision rules to determine which claims result from a single visit. In this study, ED visits were defined by the patient ID and the date of services provided. This slightly overcounts actual ED visits, as some patients spend the night in the ED, generating records of procedures performed on two separate days. The user may decide to group procedures taking place on adjacent days as belonging to the same ED visit, but this will slightly undercount actual ED visits, since some patients do visit the ED on multiple days in a row. There are no time stamps in Medicare data, so we cannot definitively address this problem.

ED visits were assembled from three Medicare administrative files: the Carrier file, the Inpatient file, and the Outpatient file. The first of these contains physician claims, while the others contain claims filed by institutions (e.g., hospitals). The type of information available in each file is different.

The Carrier file includes two sources of ED visits: the Line Place of Service code and the Line HCPCS code. The HCPCS code records the procedure for which the physician is charging Medicare. The Line Place of Service code is used in the carrier file to specify where the physician performed a procedure. Code 23 refers to the ED.

The Outpatient file also includes two sources of ED visits: the revenue center code and the HCPCS code. Revenue centers 450–459 and 981 are specific to the ED. See Table 2.3 for details. In addition, the Outpatient file includes the five ED-specific E&M codes.

The Inpatient file has less detail on ED visits because of the Medicare rules for payment for inpatient visits and related expenses. The Prospective Payment System (PPS) rules require that certain outpatient hospital services are bundled together into the single payment for an inpatient stay. Since 1991, diagnostic services pro-

Table 2.3: Key Emergency Department revenue codes

Revenue Code	Definition
0450	Emergency room—general classification
0451	Emergency room—EMTALA emergency medical screening services
0452	Emergency room—ER beyond EMTALA screening
0456	Emergency room—urgent care
0459	Emergency room—other
0981	Professional fees—Emergency room

Source: <http://www.resdac.org/cms-data/variables/revenue-center-code>

vided by the hospital up to 3 calendar days before an inpatient admission are not separately reimbursed—they are considered to be covered by the single inpatient PPS payment. Nondiagnostic outpatient services “related to” a patient’s inpatient stay are also bundled with the inpatient stay. From 1998 until 2010, “related to” was defined as having an exact 4-digit match on the ICD-9 diagnosis code assigned to the inpatient stay and to the outpatient services. Starting June 25, 2010, all outpatient nondiagnostic services are considered “related to” the inpatient stay, except for ambulance services and some renal dialysis services. For much more detail on this policy, see the Medicare Claims Processing Manual, Chapter 3, section 40.3.⁴ The result of this rule is that hospitals do not submit bills to Medicare that outline each procedure performed as part of the stay (or in the three days before, where relevant). They do, however, provide limited information in the form of total charges by revenue center. We can identify emergency department visits related to an inpatient stay by looking for charges to the ED revenue centers from Table 2.3. Note that physicians may separately charge for their evaluation and management services in some circumstances,

4. Some hospitals are not subject to the 3-day window provision, including critical access hospitals, psychiatric hospitals/units, inpatient rehab facilities/units, long-term care hospitals, children’s hospitals, and cancer hospitals. These facilities are instead subject to a 1-day window.

even though the hospital must wrap its ED charges into the inpatient stay. Thus, for some visits (22% of all ED visits identified in the sample), we see physician charges in the Carrier file and institutional charges in the Inpatient file.

In all, 73% of ED visits identified in the sample had both an institutional ED source (either inpatient or outpatient) and a physician ED source (i.e., from the carrier file) for an individual patient on a single day. An additional 5% had both a physician ED charge and an inpatient admission with ED charges within 3 days.

Generating a list of ED indicator procedures

To create a list of procedures that indicate that a visit with a moderate level E&M code (99282–3) was appropriate for the ED, we combined empirical analysis of claims data and physician review. After creating a list of procedures associated with ED visits, we created a list of procedures frequently performed in primary care visits. Primary care visits were identified using a list of E&M codes for primary care visits as defined in Chang et al. (2011). All procedures performed on the same day as an identified primary care visit were captured across both the carrier and outpatient files to ensure that lab tests and imaging studies performed outside the primary care doctor’s office were included in the sample.

A list of potential ED indicator procedures to distinguish among moderate severity ED visits was compiled of procedures frequently performed in the ED that were not frequently performed in primary care visits. These codes were reviewed by an emergency medicine physician (MFB), a geriatrician (RLK), and an urgent care physician for clinical logic—whether the procedures seem to suggest that the patient required care in an ED. The final list contains 120 procedure codes. See appendix A.

Outcomes

Two outcomes were used to validate the new algorithm in comparison to the Ballard version of the Billings/NYU algorithm: death (at 1 week and 1 month following the ED visit) and hospitalization (at 1 day and 1 week following the ED visit). These outcomes have the advantage of being available in the data set used for this study. They are not ideal for the purpose of discriminating between ED visits that could have been primary care treatable and those that are not, since they are quite serious outcomes. Nevertheless, these outcomes are useful for validating that visits identified as severe are indeed associated with either severity of illness (resulting in death) or need for more intensive care than can be provided in an outpatient setting (resulting in hospitalization).

Because the new algorithm uses hospitalization as a criterion in categorizing ED visits, an alternative version of the new algorithm was devised that uses only the E&M criteria. ED visits that resulted in an inpatient admission but have no record of an ED-specific E&M code are left uncategorized with this version of the new algorithm. These uncategorized visits represent 6–7% of ED visits.⁵ This exercise should be considered to validate the approach of using E&M criteria with the ED indicator procedure list to classify ED visits, rather than the new measure *per se*.

Validation sample

Validation analyses were performed on a subset of observed ED visits. Only beneficiaries who were covered by FFS Medicare parts A and B for all possible months of the year were included in the analysis, to ensure we had as much information as possible on medical care received. In general, beneficiaries were required to have 12

5. Visits uncategorized by the alternative version of the algorithm but categorized by the main version of the algorithm represent 6.8% of all ED visits identified in the sample and 6.2% of ED visits by people who were covered by FFS Medicare parts A and B for all possible months of the year, as defined for the validation sample below.

months of part A, 12 months of part B, and 0 months of HMO coverage. However, this rule would exclude people who died or who aged in to the program, resulting in partial years of coverage. To capture these valid partial years of coverage, two additional groups of beneficiaries were retained in the analysis: both groups had matching numbers of months (fewer than 12) of A and B coverage and no HMO coverage during the year. The aged-in group were 64 years old at the beginning of the year, while the group of beneficiaries who died had parts A and B coverage termination codes indicating death.

Some of the death information included in the denominator file was clearly in error. For example, approximately 22,000 people have a “valid date of death” flag in both 2011 and 2012 (suggesting they had died in both years). We did not want to lose any deaths that could be kept. In some cases, the two reports of death matched (e.g., both years’ denominator files reported the same death date. These beneficiaries were retained in the analysis and the earlier, incorrect death flag was removed (in all of these cases, the death date reported in both denominator files was in 2012). In cases where two death dates were provided and the dates did not match, the beneficiaries were removed from the validation analysis.⁶ Any ED visit that takes place after the stated date of death is also removed from the analyses.

Statistical methods

Generalized estimating equations were used to estimate the odds ratio of death or hospitalization within the specified time frame associated with a more severely rated ED visit compared with a primary-care treatable visit. The method accounts for correlation due to potentially repeated observations for each beneficiary with any ED visits observed over the two years of data. Deaths have been populated in the dataset

6. No ED visits were removed from the visit counts of all ED visits (N=2,644,545) for bad death dates or for not meeting coverage criteria.

through late March of 2013 and appear to be approximately equally complete over all months of data.

The model estimated is as follows:

$$\text{logit}(\Pr(o_{is} = 1|\mathbf{x}_i)) = \boldsymbol{\beta}\mathbf{x}_i + \gamma\text{AlgCategory}_{is} \quad (2.1)$$

$$o_{is} \sim \text{Bernoulli}(\pi_{is}) \quad (2.2)$$

where o_{is} is a binary variable for whether beneficiary i died (or was hospitalized) during the relevant time period after ED visit s , \mathbf{x}_i is a vector of person-specific explanatory variables, and AlgCategory_{is} is a categorical variable indicating which algorithm category was assigned to the ED visit. The omitted reference category was “nonemergent” for the Ballard algorithm and “primary care treatable” for the E&M algorithm. An independence working correlation matrix was specified, as was the Huber-White robust variance-covariance estimator.

Included covariates were generally available in the data for each beneficiary by year. For example, the comorbidity measure—Hierarchical Condition Category (HCC)—categorization was run on 2011 and 2012 data separately. Race and sex variables were reported in the denominator file by year and wouldn’t necessarily match across the two years of data. Home state and county also came from the denominator file, so different values are possible in 2011 vs. 2012. The covariates included in the model are from the year in which the ED visit took place.

A series of binary variables representing the 2014 HCC diagnosis groups was included in each model to control for beneficiaries’ varied levels of underlying health. These are the conditions used by CMS to risk adjust Medicare Advantage payments.

Table 2.4: Evaluation and Management algorithm: categorization of Medicare sample ED visits

		Evaluation and Management algorithm categories				
		Primary Care Treatable	ED Care Needed	Total Classified	Unclassified	Total
Total	N	462,191	2,010,083	2,472,274	172,271	2,644,545
	percent	17.48%	76.01%	93.49%	6.51%	100.00%

Note: includes all ED visits identified in sample (2011–12 LDS 5% Medicare sample).

Results

Classification of visits

The new algorithm categorizes 93.5% of all ED visits identified in this sample.⁷ Unclassified visits did not result in an inpatient admission and also did not include an evaluation and management (E&M) code from the list of 7 codes that are part of the algorithm. Just over 76% of visits were coded as requiring ED care.

A small number of unclassified visits (12,673=0.5% of the total visits) had a Type B ED E&M code. Type B Emergency Departments do not meet the Medicare definition of a Type A ED, but are still subject to EMTALA (CMS 2012b, Ch. 4, 160). The Type B E&M codes follow the same guidelines as the standard ED E&M codes. However, the Type B EDs may have fewer resources or be open fewer hours than a standard ED. These visits were thus left unclassified, rather than trying to fit them in with the standard ED visits. Other researchers may choose to include them in the classification, which would be easy to do.

Visits can be classified as requiring ED care in three ways: they could result in an inpatient admission, they could have one of the most severe E&M codes (critical

7. The distribution is not substantially different in 2011 versus 2012—the proportions in each category are generally less than one percentage point different—so the two years are presented together throughout this section.

care, 99284, or 99285), or they could have an intermediate severity E&M code plus one of the ED indicator procedures. In fact, just 1.4 percent of ED visits were coded as requiring ED care using the ED indicator procedure criterion. Another way of looking at this figure is that of all ED visits with an intermediate severity E&M code (99282 or 99283), 7.5 percent had an ED indicator procedure. One interpretation is that the ED indicators are not a very important part of the algorithm. However, this may be an artifact of the special population used in these analyses. It would be important to compare results in a more general ED visit dataset that includes all ages and payers.

Comparison to Ballard/Billings algorithm classifications

Ballard et al. (2010) specification

Table 2.5 compares the results of the E&M algorithm to the results of the Ballard algorithm classifying the total sample of ED visits. See appendix B for a full tabulation of the ED visit sample using the original specification of the Billings/NYU algorithm, as well as a comparison of the Ballard specification and the original specification of the Billings/NYU algorithm.

Table 2.5 includes counts of visits and row, column, and total percentages. Row percentages are the distribution of Ballard categories for a single E&M algorithm category. For example, 59.0% of visits rated “primary care treatable” by the new algorithm are categorized as non-emergent by the Ballard algorithm, 4.9% are rated intermediate by the Ballard algorithm, and 29.9% are rated “emergent” by the Ballard algorithm. The remaining 6.2% are not fully classified by the Ballard algorithm. Column percentages in Table 2.5 give the distribution of E&M algorithm categories for a single Ballard algorithm category; for example, 35.9% of visits rated “nonemergent” by the Ballard algorithm are rated “primary care treatable” by the E&M algorithm.

Total percentages give the cell percentage of all ED visits identified in the sample; for example, 10.3% of all visits are rated “primary care treatable” by the new algorithm and “nonemergent” by the Ballard algorithm.

Table 2.5: Comparison of categorization of Medicare sample ED visits: E&M algorithm and Ballard algorithm

E&M category		Ballard category					Total
		Nonemergent	Intermediate	Emergent	Special category	Unclassified only	
Primary care treatable	n	272,783	22,843	138,033	7,397	21,135	462,191
	row pc	59.0%	4.9%	29.9%	1.6%	4.6%	100.0%
	col pc	35.9%	23.7%	8.1%	37.8%	32.8%	17.5%
	total pc	10.3%	0.9%	5.2%	0.3%	0.8%	17.5%
ED care needed	n	422,058	69,420	1,493,476	9,076	16,053	2,010,083
	row pc	21.0%	3.5%	74.3%	0.5%	0.8%	100.0%
	col pc	55.6%	71.9%	87.6%	46.4%	24.9%	76.0%
	total pc	16.0%	2.6%	56.5%	0.3%	0.6%	76.0%
Unclassified	n	64,952	4,290	72,749	3,092	27,188	172,271
	row pc	37.7%	2.5%	42.2%	1.8%	15.8%	100.0%
	col pc	8.6%	4.4%	4.3%	15.8%	42.2%	6.5%
	total pc	2.5%	0.2%	2.8%	0.1%	1.0%	6.5%
Total	n	759,793	96,553	1,704,258	19,565	64,376	2,644,545
	row pc	28.7%	3.7%	64.4%	0.7%	2.4%	100.0%
	col pc	28.7%	3.7%	64.4%	0.7%	2.4%	100.0%

Note: pc=percentage

We consider a visit to have concordant ratings from the two algorithms if the new algorithm rating is “primary care treatable” and the Ballard rating is “nonemergent” or if the new algorithm rating is “ED care needed” and the Ballard algorithm rating is emergent. In all, two-thirds of ED visits have concordant ratings from the two algorithms: 10.3 percent are rated “primary care treatable” and nonemergent, while 56.5 percent are rated “ED care needed” and emergent.

Over one-fifth of the ED visits have discordant ratings from the two algorithms: 5.2 percent of the total are rated “primary care treatable” by the new algorithm and “emergent” by the Ballard algorithm—for these visits, the Ballard algorithm gave a more severe rating. In 16.0 percent of all ED visits, the opposite was true: they were rated “ED care needed” by the new algorithm and “nonemergent” by the Ballard algorithm, with the new algorithm rating the visits as more severe than the Ballard algorithm.

In one-quarter of the discordant visits, the Ballard algorithm rated the visit more severe than the E&M algorithm. The five most common primary diagnoses for these “Ballard-higher” visits were chest pain NOS (not otherwise specified), shortness of breath, lumbago, sciatica, and sprain of the shoulder/arm NOS. Some of these diagnoses seem quite severe—in particular, chest pain and shortness of breath. Unfortunately, we don’t have complete records of these visits, so we can’t determine why the physicians who saw these patients did a very modest workup when the primary diagnosis seems potentially severe. We do know that in 93 percent of these discordant visits with chest pain or shortness of breath as the primary diagnosis, the ED-specific E&M code was 99282 or 99283, so the rating of the visit as “primary care treatable” was based on the lack of any ED indicator procedures. So in at least 93 percent of these visits, the patient was discharged without receiving lab tests of blood gases, creatine kinase, troponin, myoglobin, or D-dimer. This may suggest that the physicians attending these patients did not have a high suspicion of heart attack or severe

hypoxemia.

The most common primary diagnoses for discordant visits where the Ballard algorithm was less severe than the E&M algorithm were headache, abdominal pain of unspecified site, abdominal pain of other specified site, dizziness and giddiness, urinary tract infection NOS, and malaise and fatigue NEC. This list of diagnoses is a good example of why it is important to take age into account when determining the severity of a diagnosis. Many diseases can have atypical presentations in elderly people. As people age, they experience a gradual loss of function that is most apparent in the response to stress (Kane et al. 2013, p. 5). Baseline function may be very similar to younger persons'; but when an organ system is stressed, the elderly person may be unable to handle that stress, which may result in atypical symptoms (or a lack of symptoms) despite serious illness. Thus the non-specific symptoms of headache, dizziness, or malaise may merit a more in-depth work-up in an elderly person than a younger one.

Validation of algorithm

Hospitalization outcomes

The predictive validity of the algorithm was tested using individual outcomes. A modified version of the E&M algorithm was created to compare the two algorithms on hospitalization outcomes. The original version of the algorithm could not be used because hospitalization is part of the criteria for categorization. The alternative specification uses only CPT codes for categorization. ED visits that appear in the inpatient file but do not have any ED-specific E&M codes are unclassified, but separated into their own category to avoid mixing them with visits that had neither an E&M code nor an inpatient admission (presumably a very different group of visits). Results are presented in Table 2.6.

Table 2.6: Predictive validity of modified E&M Algorithm and Ballard algorithm: hospitalization within 1 day and 1 week from ED visit

Outcome:	(1)	(2)	(3)	(4)
Hospitalization	1 day	1 day	1 week	1 week
E&M Alg: PCT ref.	1		1	
	[1,1]		[1,1]	
ED Care needed	11.70		7.681	
	[11.49,11.91]		[7.566,7.797]	
Unclassified	1.519		1.372	
	[1.473,1.566]		[1.336,1.409]	
Unclassified: inpatient only	121.7		78.09	
	[117.0,126.6]		[74.93,81.39]	
Ballard Alg: Minor ref.		1		1
		[1,1]		[1,1]
Emergent		6.717		5.259
		[6.647,6.788]		[5.208,5.311]
Intermediate		2.170		1.855
		[2.124,2.217]		[1.818,1.892]
Special category		2.183		2.100
		[2.075,2.296]		[2.002,2.203]
Unclassified		0.678		0.740
		[0.654,0.704]		[0.716,0.765]
Observations	2448112	2448112	2448112	2448112

Exponentiated coefficients; 95% confidence intervals in brackets

All ORs significant at $p < 0.001$; all models adjusted for HCC score, race, sex, and age.

Columns 1 and 2 of Table 2.6 display the results of the analysis of hospitalization within one day of an ED visit from the new method (column 1) and the Ballard method (column 2). Results are presented as odds ratios, with the reference category being the least severe category for the algorithms: for the E&M algorithm, it's the "primary care treatable" category, while for the Ballard algorithm, it's the "nonemergent" category.

Both algorithms are good predictors of hospitalization within one day of an ED visit. The odds ratio of 11.70 in column 1 indicates that the odds of being hospitalized within one day of an ED visit are 11.70 higher when a visit is rated "ED care needed" by the new algorithm compared to visits rated "primary care treatable" by the new algorithm. Similarly, the odds ratio of 6.72 in column 2 indicates that the odds of hospitalization within one day of an ED visit are 6.72 times higher for visits rated "emergent" by the Ballard algorithm, compared to visits rated "nonemergent" by the

Ballard algorithm. The confidence intervals for the odds ratios from the two models do not overlap, suggesting that they are statistically significantly different from each other. The larger odds ratio for the new algorithm compared to the odds ratio for the Ballard algorithm suggests that the new algorithm is better at discriminating between more and less severe ED visits.

Results for hospitalization within one week of an ED visit (columns 3 and 4 in Table 2.6) are similar to those for hospitalization within one day of a visit; again, both algorithms are good predictors of hospitalization. The odds ratio for a visit rated “ED care needed” by the new algorithm versus one rated “primary care treatable” is 7.68, indicating a person with a more severely rated visit has more than 7 times the odds of hospitalization within a week of the visit, compared to a person with a less severely rated visit. The parallel result for the Ballard algorithm is an odds ratio of 5.26 for a visit rated “emergent” versus one rated nonemergent. The confidence intervals for the estimated odds ratios for the two algorithms do not overlap, again suggesting a statistically significant difference in the two algorithms’ discrimination between more and less severe ED visits.

Death outcomes

The predictive validity of the algorithm was also tested using death at 1 week and 1 month after each ED visit as the outcome of interest. Results are compared to the Ballard version of the Billings/NYU algorithm in Table 2.7.

Both algorithms are good predictors of death within one week and one month of an ED visit, with odds ratios larger than 1 for more severe visits compared to less severe visits. Columns 1 and 2 of Table 2.7 contain the results for the analysis of death within one week of an ED visit. After adjusting for comorbidities, race, sex, and age, a person with a visit rated “ED care needed” by the new algorithm faces 3.70 times the odds of death within 1 week of the visit compared to a person with a visit rated

Table 2.7: Predictive validity of E&M Algorithm and Ballard algorithm: death at 1 week and 1 month from ED visit

Outcome:	(1)	(2)	(3)	(4)
Death	1 week	1 week	1 month	1 month
E&M Alg: PCT ref.	1		1	
	[1,1]		[1,1]	
ED Care needed	3.704		3.084	
	[3.523,3.895]		[2.986,3.186]	
Unclassified	3.450		2.969	
	[3.214,3.704]		[2.814,3.133]	
Ballard alg: Minor ref.		1		1
		[1,1]		[1,1]
Emergent		5.164		3.032
		[4.946,5.391]		[2.959,3.106]
Intermediate		1.541		1.566
		[1.408,1.687]		[1.492,1.644]
Special category		2.418		1.808
		[2.059,2.839]		[1.638,1.995]
Unclassified		4.389		2.786
		[4.067,4.737]		[2.638,2.942]
Observations	2448112	2448112	2448112	2448112

Exponentiated coefficients; 95% confidence intervals in brackets

All ORs significant at $p < 0.001$; all models adjusted for HCC score, race, sex, and age.

“primary care treatable.” Using the Ballard algorithm to categorize visits, we find that a person with a visit rated “emergent” has 5.16 times the odds of death within a week of the visit compared to a person with a visit rated “nonemergent.” The OR for a more severe visit versus a less severe one is larger for the Ballard algorithm than the new algorithm, indicating that the Ballard algorithm might be a better predictor of death within a week of a visit than the new algorithm.

The analysis of death within one month of an ED visit shows no difference in performance between the two algorithms. The odds of death within one month of an ED visit are 3.08 times higher for a person with a visit rated “ED care needed” compared to one rated “primary care treatable” by the new algorithm, and the odds are 3.03 times higher for a person with a visit rated “emergent” compared to one rated “nonemergent” by the Ballard algorithm.

Comparison of outcomes in visits with discrepant algorithm categorizations

Since the algorithms disagree on such a large proportion of visits, we analyzed the subset of visits with discrepant algorithm results. Specifically, we estimated the odds of death within 1 week or hospitalization within one day of an ED visit for those visits where either:

1. the Ballard algorithm rated “emergent” and the new algorithm rated “primary care treatable” (ballard_higher=1, B=E/E&M=PCT), or
2. the Ballard algorithm rated “nonemergent” and the new algorithm rated “ED care needed” (ballard_higher=0, B=NE/E&M=EDCN).

Visits where the algorithms agree in their ratings were excluded from this analysis.

Table 2.8: Predictive validity of modified E&M Algorithm and Ballard algorithm when the algorithms disagree: death within 1 week and 1 month of ED visit

	(1) Hospitalized next day	(2) Death 1 week	(3) Death 1 week excluding Cardiac Arrest
ballard_higher=1	0.722 [0.706,0.739]	2.036 [1.881,2.205]	1.156 [1.050,1.273]
Observations	522881	518135	517678

Exponentiated coefficients; 95% confidence intervals in brackets

Model 1 uses the modified version of the new algorithm that drops the inpatient criterion.

All ORs significant at $p < 0.001$; all models adjusted for HCC score, race, sex, and age.

The discrepant visit analysis encompasses approximately 320,000 beneficiaries contributing an average of 1.6 visits each.⁸ Results are presented in Table 2.8. As

8. The hospitalization analysis and death analyses use slightly different samples because the modified version of the algorithm used in the hospitalization analysis creates more discrepant results. This happens when a visit that resulted in an inpatient visit had ED-specific E&M codes in the Carrier file of physician claims that do not result in an ED-care needed rating. The analyses were repeated using the exact same set of visits (that is, only the visits that were discrepant in both versions of the E&M algorithm. A further 187 visits with hcc110=1 [cystic fibrosis] were also excluded

we would expect, the results confirm that the modified version of the new algorithm (without the inpatient criterion) is a better predictor of hospitalization within one day than the Ballard algorithm. The binary predictor variable of interest is “ballard_higher,” which takes the value 1 when the Ballard algorithm result was “emergent” and the new algorithm result was “primary care treatable” (B=E/E=PCT); it takes the value 0 when the Ballard algorithm result was “nonemergent” and the new algorithm result was “ED care needed” (B=NE/E&M=EDCN). Coding the variable this way allows us to analyze both algorithms’ results in the same model. Column 1 of Table 2.8 tells us that the odds ratio on ballard_higher was 0.72, which is less than 1. This indicates that visits coded B=E/E&M=PCT were less likely to result in hospitalization within one day than visits coded B=NE/E&M=EDCN. In other words, when the two algorithms differ in their ratings, the visits rated “primary care treatable” by the E&M algorithm are less serious (as measured by their odds of hospitalization) than the visits rated “nonemergent” by the Ballard algorithm. This finding corresponds to columns 1 and 2 in Table 2.6 where the odds ratio on the new algorithm “ED care needed” vs “primary care treatable” was larger than that on Ballard “emergent” vs “nonemergent.”⁹

The death within one week analysis shows the opposite (column 2 of Table 2.8): the odds ratio is 2.04 on the ballard_higher predictor variable, confirming the finding in Table 2.7 columns 1 and 2 that the Ballard algorithm is a better predictor of death within 1 week than the new algorithm.

because these visits perfectly predict death and are therefore dropped from the death analysis) as a sensitivity analysis. The OR on ballard_higher for death within 7 days changes very little (from 2.04 to 2.26). The OR on ballard_higher for hospitalization changes a lot, from 0.72 to 0.25. This is to be expected, since we are excluding visits that resulted in a hospitalization, but that had minor E&M codes. This causes these visits to be classified as “primary care treatable” when the hospital admission is not taken into account. Excluding these visits improves the apparent performance of the new algorithm.

9. These findings must correspond in this way, since all the “action” in the analysis in Table 2.6 must be coming from visits where the two algorithms disagree.

However, this finding is driven by a single diagnosis code: 427.5 Cardiac arrest. There are 457 eligible discrepantly coded ED visits with a primary diagnosis code of cardiac arrest where the new algorithm categorizes the visit as primary care treatable—that is, the evaluation and management code assigned for the visit was 99281 or the E&M code was 99282 or 99283 and none of the ED indicator procedures also appeared in the claims related to the visit; in 450 of those 457 cases, the person died within a day of the visit. The ED indicator procedure list includes CPR and cardioversion, which we would expect to see in cases of cardiac arrest; that we do not see these procedures suggests that the patients may have arrived in the ED dead. Excluding ED visits with cardiac arrest as the primary diagnosis code reduces the size of the effect from an odds ratio of 2.04 to 1.16 for death within 1 week when the Ballard algorithm gives the more severe result (B=E/E&M=PCT) versus when the E&M algorithm gives the more severe result (B=NE/E&M=EDCN). See column 3 of Table 2.8.

When the death analyses are performed on the full dataset (not just visits with discrepant algorithm results) excluding 7,114 visits with a primary diagnosis of cardiac arrest, the Ballard algorithm is no longer a better predictor of death than the new algorithm. See Table 2.9. The odds ratios for death within one week of the ED visit are statistically equal across the two algorithms (the 95% confidence interval for “ED care needed” overlaps the point estimate for Ballard Emergent). The odds ratio for death within one month is slightly higher for the new algorithm.

Table 2.9: Predictive validity of standard and modified E&M Algorithm and Ballard algorithm excluding cardiac arrest: death within 1 week and 1 month of ED visit

Outcome:	(1)	(2)	(3)	(4)
Death	1 week	1 week	1 month	1 month
E&M Alg: PCT ref.	1		1	
	[1,1]		[1,1]	
ED Care needed	4.134		3.144	
	[3.900,4.381]		[3.038,3.253]	
Unclassified	4.275		3.161	
	[3.954,4.623]		[2.990,3.342]	
Ballard alg: Minor ref.		1		1
		[1,1]		[1,1]
Emergent		4.182		2.766
		[4.005,4.366]		[2.700,2.834]
Intermediate		1.508		1.553
		[1.377,1.651]		[1.479,1.630]
Special category		2.378		1.794
		[2.024,2.793]		[1.625,1.981]
Unclassified		4.260		2.751
		[3.945,4.600]		[2.604,2.907]
Observations	2440998	2440998	2440998	2440998

Exponentiated coefficients; 95% confidence intervals in brackets

All ORs significant at $p < 0.001$; all models adjusted for HCC score, race, sex, and age.

Discussion

We developed and validated a new algorithm to determine the primary care treatability of ED visits using Medicare administrative data. Our new algorithm measures healthcare utilization using the highest E&M code and procedures performed during the ED visit. It proposes a new measure of primary care treatability of ED visits that solves some of the problems presented in applying the Billings/NYU algorithm.

To validate our model we compared the new measure's performance to the established Ballard variant of the Billings/NYU algorithm in predicting hospitalization and death following an ED visit. Because our algorithm included hospitalization in the original phase, we modified the model to be able to compare it to the Ballard version of the Billings/NYU algorithm. Both the new algorithm based on procedure codes and the Ballard version of the Billings/NYU algorithm were able to predict hospitalizations and deaths after ED visits in a Medicare FFS population. The new algorithm is a somewhat better predictor of hospitalization after an ED visit and a slightly worse predictor of death after an ED visit than the Ballard algorithm in analyses that controlled for race, sex, and HCC comorbidities.

The Ballard algorithm's superior ability to predict death in this sample is driven by visits with a primary diagnosis of cardiac arrest, some of which are rated "primary care treatable" by the new algorithm. These visits may include patients who are dead on arrival to the ED. Depending on the state, a physician may be required to pronounce death, which could result in a low-severity E&M code. In these cases, an algorithm based on procedural codes will not capture the severity of the patient's condition.

Researchers wishing to classify the severity of ED visits observed in claims data will want to consider several factors in choosing between the two algorithms.

The E&M algorithm uses procedure codes to determine whether an ED visit was

primary-care treatable, while the Ballard algorithm uses diagnosis codes. The advantage to the procedure code approach is that it captures the medical care deemed appropriate by a physician who actually saw the patient.

Older people are more likely to have an increased number of comorbidities, and an increased number of comorbidities will worsen the severity of the illness; for example, a young patient with an urinary tract infection can be managed as outpatient with oral antibiotics, but an elderly patient with diabetes and renal failure, might need to be admitted to the hospital for intravenous antibiotics when presenting with the same urinary tract infection. Both increasing age and comorbidities are associated with increased mortality (Chen et al. 2012; Lundgren et al. 2009; Piccirillo et al. 2008; Rich 2005; Yancik et al. 1998).

The diagnosis-based approach of the Ballard algorithm avoids problem of basing the rating of a visit on the amount of care received, with its associated endogeneity issues. However, diagnoses may be little more objective than procedures. Different physicians and medical billing coders may apply different ICD9 codes to the same patient, depending on their approach to taking a medical history, their clinical experience, the patient's comorbidities, and so on. Song et al. (2010) showed that moving from a low-intensity practice area to a high-intensity area doubled the number of diagnoses recorded for Medicare beneficiaries.¹⁰ In some cases, diagnosis coding can also affect physicians' financial incentives; for example, some lab tests are only reimbursed if a particular diagnosis is present. HbA1c tests are considered by Medicare to be appropriate only for management of diabetes, not as a diagnostic test (CMS 2011, section 190.21)

Investigators and clinicians who wish to evaluate the ED the ED healthcare utilization of a particular disease or condition—for example, patients with congestive

10. The issue is more profound for Medicare Advantage beneficiaries. See Geruso and Layton (2014) on “manipulable medical coding” to increase revenue in the context of risk adjusted payments.

heart failure or cardiovascular disease—will benefit from using the new algorithm. The Billings/Ballard algorithms do not allow visits to vary in severity within a diagnosis, limiting their use in these questions.

Another significant advantage of our new proposed model is that the implementation of ICD10 will require a remapping of the Billings/NYU/Ballard algorithm diagnoses, which are coded in ICD9, while the E&M algorithm will not be affected by this change.

There is lack of agreement on how to judge inappropriate or primary care treatable populations. The use of expert opinion, self-rating by patients, review of department activities and subsequent admissions have all failed to determine appropriateness when applied to the patient level (Bezzina et al. 2005). When evaluating the appropriateness of an ED visit one must consider variables that are unable to be coded/found on claims data like the time of the day and day of the week. Access to a primary care provider, convenience, uninsurance and poverty are reasons why many patients with primary care treatable diseases present to an ED, despite the waiting times and the increased cost of an ED visit (Kellermann and Weinick 2012; Weinick et al. 2010; Begley et al. 2006; Pitts et al. 2010) Studies that evaluated the access to primary care in children with Medicaid insurance found a strong temporal relation between decreased visits to the ED and increased access to primary care services (Piehl et al. 2000). Specific services that may be responsible for the decreased ED use include the expanded availability of primary care physicians and the use of telephone triage systems (Piehl et al. 2000). Improving primary care access and scope of services may reduce ED use. Focusing on systems issues rather than patient characteristics may be a more productive strategy to improve appropriate use of emergency medical care (Lowe et al. 2005).

Limitations

The ideal validation of the new algorithm would be to compare it to a gold standard. In this context, that would be a set of ED visits rated by physicians with access to the full medical record of the visit. Hospitalizations and mortality are not perfect proxies for that gold standard. We would expect that some people who need to be seen in the ED could be treated without necessitating a hospital admissions, for example. Indeed, the Billings algorithm was created with a sample of ED visits that specifically did not result in an admission. Similarly, a medical issue need not carry the threat of death in order to be an appropriate reason for an ED visit. The new algorithm was designed to pick up indicators in claims data that a visit required ED care by looking at the amount of healthcare resources that were used, how intensively the ED physician evaluated the patient, and at whether some procedures were performed that would be difficult to access in a primary care setting. These design features provide some face validity for cases where hospitalization and death are not convincing signs of the necessity of ED care.

The new algorithm relies on claims data to determine which procedures were performed during an ED visit. Some researchers and journalists argue that there is significant “upcoding” in Medicare billing, particularly in ED visits (see, e.g., Brunt (2011) and Schulte and Donald (2012) on ED visit trends). They argue that lax Medicare oversight has resulted in rampant overcharging for a higher level E&M code than is merited by the work actually done. This may not be an accurate depiction of the scope of the problem. Coding rules are complicated; coding experts facing no financial incentives for upcoding show poor to fair agreement on the appropriate E&M code to apply to an ED visit (Bentley et al. 2002).

There are multiple audit mechanisms to investigate potential overcharging in Medicare billing. The limited publicly available results differ in their assessments of

the impact of upcoding in Medicare ED claims. The Office of the Inspector General of the Department of Health and Human Services completed an audit of evaluation and management charges that included a medical review of charts related to each claim included in their sample (OEI 2014). That study found a large proportion of E&M claims were coded incorrectly: 42.4%, of which three-fifths were upcoded and two-fifths were downcoded. Nearly 80% of incorrectly coded claims were within one level of the “correct” code determined by the audit.

By contrast, the annual CERT (Comprehensive Error Rate Testing) audit program sponsored by CMS finds a much lower error rate: across all E&M codes, the error rate is 13.9% (measured in dollars, not claims); for ED E&M codes the error rate is 7.1%, of which 76% is due to errors in coding (CMS 2012c).¹¹ About 70% of the error rate is overpayment, the remaining 30% is underpayment.¹²

Assuming significant upcoding does exist, the design of the E&M algorithm makes it relatively robust to miscoding of a single level, which seems to account for the bulk of coding errors. A visit with a “true” E&M level of 99281 (least severe visit) could only be bumped from the “primary care treatable” category to the “ED care needed” category by being upcoded 3 levels to 99284 (second most severe visit) or by being upcoded 1 or 2 levels and having an additional charge for an ED indicator procedure. Less than 6% of visits with an E&M code of 99283 or lower have one of these ED indicator procedures.

A simulation of the effect of misclassification of the magnitude suggested by the OIG audit report suggests that if the visits had been classified according to their “true” E&M code, the proportion of visits rated “primary care treatable” would have gone from 18.1% to 19.7% (2.5–97.5 percentile range based on 1000 replications:

11. Insufficient documentation accounts for 29.1% of the ED E&M errors, and “other” error for 1.6%. Medical necessity and no documentation each account for less than 0.1% of ED E&M error.

12. It is not clear why the error rates are so different for the two audit reports. The OIG report specifically notes a few differences in sampling and purpose between the two audits, though it seems unlikely that these differences would account for such a large discrepancy.

19.66% to 19.72%), while the proportion rated “ED care needed” would have gone from 75.4% to 73.8% (2.5–97.5 percentile range based on 1000 replications: 73.77% to 73.83%).

Perhaps more saliently, the new algorithm has been validated using the actual E&M codes charged and paid—and performs quite well. Over time, if coding practices shift significantly, the algorithm may need to be validated again.

Conclusions

This study introduces a new claims-computable measure of the primary care treatability of ED visits. The new algorithm shows good validity. It is a strong predictor of both hospitalizations and deaths, with performance similar to or better than the most commonly used existing algorithm to assess the severity of ED visits.

The procedure-based approach of the new algorithm allows researchers to use the clinical judgment of the ED physician who saw the patient to determine the likely severity of each visit, rather than depending on a calculation of average visit severity for a potentially very different population. The new algorithm may thus provide a useful tool for investigating ED use in Medicare beneficiaries.

Chapter 3

Paper Two: Increased primary care visits associated with decrease in primary care treatable ED visits in Medicare population

Introduction

The emergency department (ED) can be a difficult place to receive primary care, particularly for elderly people. Older people are more likely to have multiple morbidities (Barnett et al. 2012), more likely to be frail (Collard et al. 2012), and more likely to use 5 or more prescriptions (NCHS 2014) than younger people. These differences make them more difficult to diagnose and appropriately treat in a setting like the ED where their health history may be less accessible and time in shorter supply than in a primary care setting (Birnbaumer 2014; Hunold et al. 2014). Nevertheless, studies have found high proportions of non-severe (Kaskie et al. 2011, 34% of visits) or primary care treatable (Pukurdpol et al. 2014, 30% of visits) ED visits among elderly people. Although elderly people may be less likely than other age groups to have a primary-care treatable ED visit, because they are heavy users of the ED, it still adds up to a large number of visits.

Several researchers have hypothesized that improving access to primary care could reduce these visits or could reduce the total number of ED visits either by providing preventive care to avoid emergencies or by providing timely access to lower acuity care. Recent observational studies investigating this link include Hunold et al. (2014), Werner et al. (2014), and D'Arcy et al. (2012). Interventions to increase primary care access, Counsell et al. (2007), Scott et al. (2004), and Coleman et al. (2001) have generally produced stronger effects than observational studies looking at naturally occurring variations in access to primary care. This difference may reflect selection effects whereby people who self-select to have a primary care relationship are sicker or otherwise more likely to visit the ED, perhaps because they are more interested in receiving any type of medical care. If such a selection process were in effect, observational studies would tend to underestimate the protective effect of a primary care relationship.

In this analysis, two approaches are used to determine whether increased use of primary care is associated with reduced rates of primary care treatable ED visits, while attempting to correct for selection effects that could bias the result. First, geographic variation in access to and use of primary care is used to estimate the effect of increased use of primary care in reducing primary care treatable ED visits. This variation is hypothesized to be due to differences in patient preferences and physician practice style that is associated with different geographic areas, but not with underlying differences in population health beyond that accounted for by observable differences in age, race, sex, and other demographic factors. Two methods to identify primary care treatable ED visits are employed. The first uses a new algorithm based on procedures performed during an ED visits, including the severity level of the Evaluation and Management (E&M) code charged for the visit. A similar analysis is used to investigate the effect of increased use of primary care on ED visits classified by the more established Billings/NYU algorithm as non-emergent. The Billings algorithm

is based on diagnosis codes associated with ED visits and has been shown in prior studies to be insensitive to changes in access to care (Lowe and Fu 2008) and ED usage patterns (Jones et al. 2011).

The second approach to estimating the effect of primary care on primary care treatable ED visits is to assess the individual-level primary care relationship, again using both algorithms. Multiple modeling methods are used to develop a plausible range of effect sizes. We also explore potential mechanisms for the observed effects by looking at weekend primary care treatable ED visits and visits for injuries.

Methods

The Minnesota ED algorithm categorizes the primary care treatability of ED visits using procedure codes. There are five procedure codes specific to ED E&M visits, ranging from a visit for a non-emergent, self-limited problem to a high-severity visit for a problem that would cause death or severe morbidity if not attended to quickly. In addition, two critical care E&M codes represent care in the ED for high-severity problems. These seven codes form the basis for the new algorithm's categorization. Most Medicare ED visits—97% of all ED visits in the sample of Medicare claims used in this study (aged, FFS Medicare beneficiaries' ED visits that included an institutional claim)—include a physician and/or institutional claim for one of these codes. Intermediate severity visits are examined further for procedures like advanced imaging studies that are generally not available in primary care physician offices to determine whether they were likely primary care treatable. Figure 3.1 presents the approach as a flowchart.

An alternative method for identifying less severe ED visits is used to provide context and a comparison: the Ballard version of the Billings NYU algorithm has been validated in a Medicare population (Ballard et al. 2010). The association between

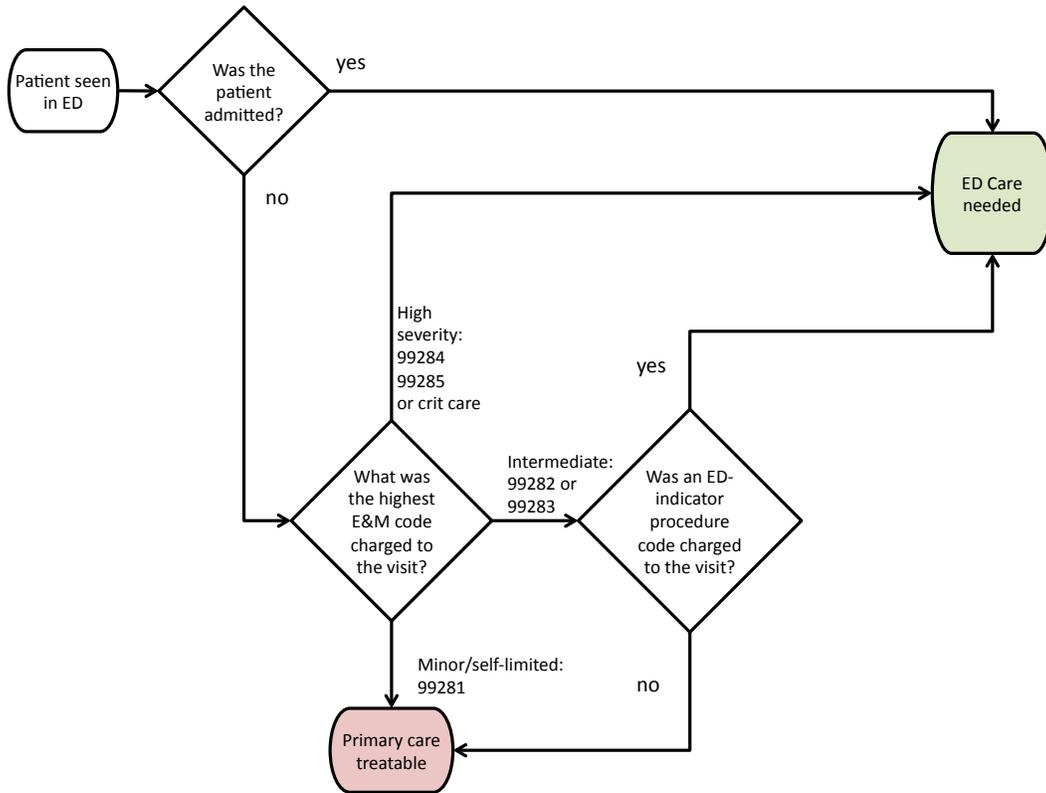


Figure 3.1: Minnesota algorithm

increased primary care use and the incidence of non-emergent ED visits as defined by the Ballard/Billings algorithm is included in the results and compared to the effect on the incidence of primary care treatable ED visits identified using the Minnesota algorithm.

Two methods are used to estimate the effect of use of primary care on the incidence of primary care treatable (PCT) ED visits: one at the hospital referral region (HRR) level and one at the beneficiary level.

HRR level analysis

The first method uses geographic variation in use of primary care to estimate the effect of increased primary care visits per capita on the number of primary care treatable ED visits observed at the hospital referral region (HRR) level. Primary care visits are defined in the Measures section below. The analysis is first carried out using the new algorithm to identify PCT ED visits, then repeated using the Ballard version of the Billings/NYU algorithm to identify non-emergent ED visits.

Negative binomial regression is used to model counts of ED visits to allow for overdispersion observed in the data. The log of number of people living in each HRR group is included in the analysis as a covariate with the coefficient constrained to 1. Thus, the exponentiated regression coefficients represent incidence rate ratios of ED visits per aged FFS beneficiary resident in the HRR group.

Individual-level analysis

The second method for exploring the relationship between receiving primary care and experiencing a primary care treatable (PCT) ED visit is specified at the beneficiary level. An unadjusted analysis of rates of PCT ED visits by whether a beneficiary has a primary care physician shows people with a primary care relationship (defined in

the Measures section below) have higher rates of PCT ED visits (9.4 visits per 100 person-months for people with a primary care relationship versus 6.7 visits per 100 person-months for those without a primary care relationship; 95% CI for difference in means 2.5 to 2.8). However, we expect this due to bias from selection effects: people who choose to have a primary care relationship might be more likely to receive any kind of health care than people without a primary care relationship. Those without a primary care relationship may simply not have any health concerns. Several analyses are used to attempt to correct for this endogeneity and give a range of estimates of the effect.

Naïve analysis: zero-inflated negative binomial, not corrected for unobserved selection

First, a naïve analysis is run, without attempting to correct for selection (beyond observables) into having a primary care relationship. A zero-inflated negative binomial (ZINB) model is used to account for the large number of zeros in the outcome variable: the number of primary care treatable ED visits experienced by a person in a year. The probability density function of the ZINB model is as follows:

$$g(y) = \begin{cases} f_1(0) + [1 - f_1(0)] f_2(0), & \text{if } y = 0, \\ [1 - f_1(0)] f_2(y) & \text{if } y \geq 1 \end{cases} \quad (3.1)$$

This allows two sources of 0-valued outcomes: one from the function $f_1(\cdot)$, a binary process (we used the logit function), and one from the function $f_2(\cdot)$, a poisson process. Non-zero valued outcomes are generated by the non-zero portion of the poisson process. The Huber/White variance estimator was used, adjusted for clustering at the person level (beneficiaries were included in the analysis for either one or two years; the mean was 1.7 years).

We compare the estimated coefficient on the primary care relationship indicator variable to the coefficients estimated in selection-corrected models for information on the success of our identification strategy and as a lower bound estimate of the treatment effect.

Endogenous treatment poisson

An endogenous treatment model for a count outcome is used. The model is estimated using the maximum-likelihood estimator from Terza (1998), as implemented in Stata's `etpoisson` command. This estimator allows for the endogeneity of treatment status in the main outcome count equation by assuming the random error terms from the treatment and outcome equations are distributed bivariate normally, conditional on the regressors in the two equations (See Terza 1998; StataCorp 2015c).

The model is as follows:

$$E(y_i | \mathbf{x}_i, d_i, \varepsilon_i) = \exp(\mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i) \quad (3.2)$$

$$d_i = \begin{cases} 1 & \text{iff } \mathbf{z}_i \boldsymbol{\alpha} + \nu_i \geq 0, \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

$$\Sigma = \begin{bmatrix} \sigma^2 & \sigma\rho \\ \sigma\rho & 1 \end{bmatrix} \quad (3.4)$$

where y_i is the count outcome; \mathbf{x}_i is a vector of explanatory variables in the outcome equation, including the individual i 's treatment status, d_i ; ε_i is a “heterogeneity term” in Terza’s parlance (or an error term); \mathbf{z}_i is a vector of explanatory variables in the treatment equation, which should include at least one variable exogenous to the outcome model; and ν_i is an error term for the treatment equation. Correlation is carried between the outcome and treatment equations by assuming that ε_i and ν_i

are distributed bivariate normally, with means of zero and the covariance matrix Σ . The variance of ν_i is normalized to 1.

A plausibly exogenous variable is included in the treatment equation, but not the outcome equation: the supply of primary care physicians per 100,000 population in the HRR group where the beneficiary resides. We would expect that living in an area with more primary care physicians would increase the likelihood that a beneficiary has a primary care relationship, but would not affect the likelihood of a primary care treatable ED visit in any other way. Unfortunately, we do not see the expected relationships between primary care physician supply and primary care relationship prevalence. A logistic regression estimating the effect of PC physician supply on prevalence of PC relationships, controlling for the total physician supply (all specialties), individual age, race, sex, and HCC score, and the rurality of the county in which the beneficiary lives gives a negative coefficient on PC physician supply. This does not mean the instrument is not valid (validity of an instrument requires correlation with the endogenous variable and orthogonality with the error term; the former is true here, though in an unexpected direction, the latter is not testable). It is possible that physician supply is correlated with some excluded variable that reduces the likelihood of primary care relationships. This would not invalidate our analysis, but because we don't know what that excluded variable is, we should be extremely cautious in interpreting our results.

Structural equation model: endogenous treatment

Stata's generalized structural equation model command `gsem` offers another way of modeling endogenous treatment (StataCorp 2015b, example 45g). The structural equation model approach includes a latent variable to carry the correlation between the treatment and the response equations (i.e., the equation determining whether a person has a primary care relationship and the equation determining how many

primary care treatable ED visits the person experiences). The treatment equation dependent variable is specified as a truncated normal variable with a lower limit of 0 if a person has a primary care relationship and an upper limit of 0 if a person has no primary care relationship. In other words, in parallel to the specification of a selection model, we imagine that treatment is a realization of a unobserved index measuring the likelihood of a person choosing to have a primary care relationship. If the level of that index is above 0, the person has a primary care relationship; if it is below 0, he/she does not. Thus, for a person with a primary care relationship, we know that the value of that unobserved index is at least 0, so we set the lower limit of the truncated normal variable to 0. Similarly, for a person without a primary care relationship, we know that the value of that unobserved index can't be as high as 0, so we set the upper limit of the truncated normal variable to 0. The latent variable L is included in both equations, with its coefficient constrained to 1 in the treatment equation. The advantage of using the `gsem` approach over the `etpoisson` command is that the former allows us to use all subjects in the treatment equation, but only the people with any ED visits in the outcome equation. In the other modeling approaches, people who have no ED visits and therefore also have no primary care treatable ED visits are not distinguishable in the outcome model from people who have at least one ED visit, but no primary care treatable ED visits: both have $y_i = 0$. In the `gsem` approach, we can set y_i as missing rather than 0 for people who have no ED visits. They will still be included in the treatment equation.

Negative binomial regression on beneficiaries with an ED visit in the previous year

We believe that people with primary care relationships may be more willing to use any type of medical care than people without primary care relationships—holding constant observable factors like age, race, and sex—due to some unmeasured positive

attitude toward medical care. For this model, we attempt to hold constant propensity to use the ED by including in the analysis only people who had any ED visit in the prior year. We then look at the effect of a primary care relationship in the current year on the primary care treatability of ED visits in the current year using a negative binomial regression with no further selection correction.

Bounded average treatment effects

The user-written Stata command `tebounds` uses multiple methods to partially identify the average treatment effect, giving upper and lower bounds for the true effect (i.e., an interval, rather than a point estimate) (McCarthy et al. 2015). The model implemented in the command allows only a binary outcome, which must be coded so that a positive outcome is coded 1; the outcome variable for this analysis is 1 if the beneficiary had *no* primary care treatable ED visits. The analysis is estimating the average causal treatment effect of a binary treatment on a binary positive outcome (that is, outcome=1 is better than outcome=0). The treatment effect is the average increase in probability of a good outcome for a person receiving the treatment compared to the same person not receiving the treatment.

$$\text{ATE} = P(Y_i = 1|T_i = 1) - P(Y_i = 1|T_i = 0) \quad (3.5)$$

Here Y_i is the outcome for person i and T_i is the indicator for treatment status. Several assumptions are imposed and relaxed to create scenarios that identify a range of plausible treatment effects. Measurement error in the treatment equation is modeled (i.e., allowing for the possibility that someone has a primary care relationship that we don't observe in the claims data).

The output from this analysis is complicated, but very rich. Two different assumptions about error in determining treatment assignment and four assumptions about

the nature of the selection process are applied in turn in several different combinations. Complete information is available in McCarthy et al. (2015), but a summary is provided here.

Treatment assignment error assumptions This analysis allows for the possibility that the observed treatment assignment can be erroneous. In this application, this would mean that we incorrectly categorize a beneficiary as having a primary care relationship that the beneficiary does not, in fact, have (a false positive) or that we miss a primary care relationship that the beneficiary has (a false negative). The analysis alternately assumes either “arbitrary” treatment assignment error (either false positives or false negatives are possible) OR “no false positives,” which assumes that only false negatives are possible.

Depending on how we conceive of the treatment effect we are estimating, either of these assumptions could be of interest. If we are narrowly defining the treatment effect as “the effect of having at least two primary care visits with the same physician over the time observed,” we could use the “no false positives” assumption. This would mean that we believe it isn’t possible to have a claim record for a primary care visit that didn’t actually take place. It would, however, allow for the possibility that we have false negatives—perhaps a person whose primary care relationship is at the VA hospital or a person whose relationship is with a clinic rather than an individual physician.

If we instead consider the true treatment effect we are estimating to be whether the beneficiary considers himself/herself to have a primary care relationship, we could be interested in the possibility that our method of identifying primary care relationships could produce false positives—people who have at least two visits with a primary care physician but do not consider that physician to be an accessible or trusted source of care. In this case, the arbitrary treatment error assumption would be of interest.

With either treatment assignment error assumption, we can provide a series of error rates between 0 and 100 percent, with treatment effects calculated assuming each level of error. For this analysis, we have looked at the effect of 0, 1, 2, 5, and 10 percent treatment assignment error.

Selection process assumptions There are five assumptions, yielding ten different scenarios about the selection process that are applied to generate treatment effect ranges; one of these assumptions (monotone treatment response) is probably not realistic in this application. The assumption is described below, but not presented in the analysis output. This reduces the assumptions applied in this analysis to four assumptions and six scenarios.

1. Exogenous selection: assumes that treatment selection (i.e., whether a person has a primary care relationship) is orthogonal to the probability of having a primary care treatable ED visit. In other words, there is no bias due to the treatment selection process
2. Worst-case selection: makes no assumptions about the potential direction of bias from the selection process. Instead, it uses the fact that the missing counterfactuals needed to calculate the true treatment effect are bounded by 0 and 1. That is, for people who had primary care relationships, we don't observe their outcome in the counterfactual case where they do not have a primary care relationship; but we do know that outcome can't be any worse than 100 percent of people having a negative outcome (i.e., a primary care treatable ED visit) or any better than 0 percent of people having a negative outcome.
3. Monotone treatment selection (MTS), with negative or positive selection: reduces the range of estimates from the worst-case selection assumption by assuming that "expected potential outcomes move in a particular direction when

comparing individuals in the treatment and control groups” (McCarthy et al. 2015). If that direction is negative, people who choose to have a primary care relationship are less likely to have a good outcome (i.e., no primary care treatable ED visits) than people who don’t choose to have a primary care relationship. If treatment selection is positive, people who choose to have a primary care relationship are *more* likely to have a good outcome than people who don’t choose a primary care relationship.

4. Monotone treatment response (MTR): assumes that people choose the treatment condition that will make them best off. In this case, we would have to believe that people want to avoid primary care treatable ED visits and that they know whether they will be more likely to avoid those visits if they have a primary care relationship or don’t have a primary care relationship. This does not seem particularly defensible in this application, so results are excluded from the output presented.
5. Monotone instrumental variable (MIV): assumes “the latent probability of a good outcome conditional on treatment assignment . . . varies (weakly) monotonically with an observed covariate” (McCarthy et al. 2015). Note that this is not as stringent a requirement as what economists would refer to as an instrumental variable (IV), and it will not be used on its own to identify the treatment effect. In this analysis, we use beneficiary age as a monotone IV. The likelihood that people have a primary care treatable ED visit goes down as age goes up because physicians are more apt to see them as frailer than younger people, and thus to give them a more thorough work-up.

The scenarios presented in the output from the bounded treatment effects analysis include exogenous selection, worst-case selection, MTS with negative and positive selection, and MTS with negative and positive selection plus MIV.

Mechanisms for a reduction in ED visits

To further explore the mechanism by which having a primary care relationship could reduce primary care treatable ED visits, two additional analyses were completed. If one of the main mechanisms for reducing primary care treatable ED visits is an availability effect, where patients choose the best care venue among those available, we would expect that patients with a primary care relationship would be more likely to have a primary care treatable ED visit when their primary care office is not available. If this is an important mechanism, we would expect to see less of a difference in the primary care treatability of weekend ED visits than in weekday ED visits. Thus, we would compare the coefficient on the primary care relationship indicator variable in the models for counts of primary care treatable ED visits using all ED visits versus the model looking only at weekend ED visits.

Because of the difficulty of teasing out selection effects from the causal preventive effect of a primary care relationship on primary care treatable ED visits, we want to know whether our identification method is successful. One possibility is to check the impact of a primary care relationship on ED visits that are extremely unlikely to be affected by primary care; if the effect is similar in direction and magnitude to the effect on primary care treatable ED visits, we would suspect that our identification strategy is unsuccessful. For this analysis, we look at ED visits for accidents unlikely to be closely related to health. These visits are identified using E-codes in the inpatient and outpatient files. E-codes are special supplementary diagnosis codes for classifying external causes of injuries. The CDC has published a categorization of ICD9 E-codes by the mechanism or cause of injury and the manner or intent by which the injury happened. Example mechanisms include cuts, drowning, falls, fire, motor vehicles, and firearms; the intent categories include unintentional, self-inflicted, assault, and undetermined (i.e., unknown whether it was accidental or on purpose). For the purposes of this analysis, ED visits with an associated E-code indicating the

intent behind the injury was unintentional or assault were counted except for visits where the cause of injury was a fall or overexertion. These exceptions were considered to be potentially highly correlated with overall health.

There is still some correlation between these injuries and overall health; for example, a very sick person may have a lower risk of being struck by a car while bicycling. The correlation between HCC score and having any ED visit with an accident as defined above is 0.09; that is, having a higher HCC score (having more comorbidities) is associated with a higher likelihood of having an ED visit with an accident. The correlation is much higher—0.20—when falls and overexertion are included with other types of accident. So excluding them makes sense, since we are trying to capture ED visits that are equally likely for people of all levels of overall health.

It is likely impossible to specify ED visits that are completely uncorrelated with health, and through that relationship correlated with whether a person has a primary care relationship, at least when working with claims data. Our measure of a primary care relationship requires having at least two primary care evaluation and management visits with the same physician during the time we observe beneficiaries. We would expect people with few health problems to be much more likely to have no doctor visits, and therefore to be classified as having no primary care relationship. As a result, we expect that the beneficiaries who meet the criteria for having a primary care relationship to be somewhat less healthy than those without a primary care relationship. This effect is probably smaller in the Medicare population than it would be in a broader population, since by definition all Medicare beneficiaries are insured and the aged fee-for-service beneficiaries in our sample will have more health problems than average.

Data

Two years of Medicare claims data (2011–2012) for a nationally representative five percent sample of Medicare beneficiaries were used. The Limited Data Set files made available by the Center for Medicare & Medicaid Services include claims for nearly all services paid for by Medicare (Part D drug claims are not available) including inpatient, outpatient, physician services, home health, durable medical equipment, hospice, and skilled nursing facilities. Information that could be used to identify individual beneficiaries has been stripped from the files according to HIPAA privacy rules. The data includes limited demographic information for all sampled beneficiaries, regardless of whether they received any medical care: beneficiary age (top-coded at 100 years of age), county of residence, gender and race, the reason for receipt of Medicare benefits (old age, end-stage renal disease, etc.), and indicators for receipt of state subsidies that serve as a proxy for dual eligibility for Medicare and Medicaid; the claim files include unencrypted physician and institution identifiers, procedure codes, and diagnosis codes, as well as submitted and paid charges. See CMS 2014 for further details.

The cohort of interest was beneficiaries aged 65 and older who had both part A and B coverage and no months of HMO coverage. Only people with a valid county-of-residence in the denominator file could be included. HRRs are defined for the fifty US states plus Washington, DC. Residents of US territories and foreign countries are excluded.

HRR assignment based on county of residence

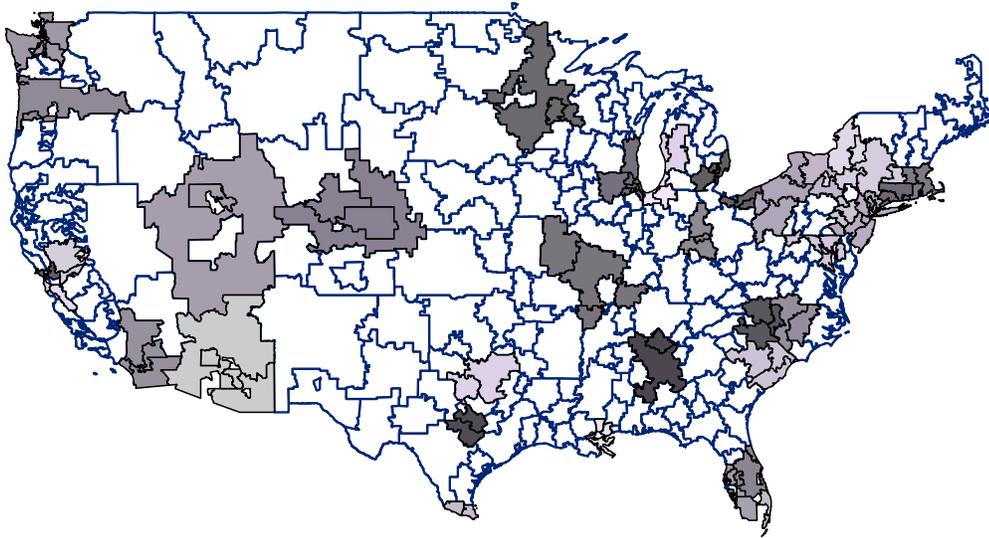
To protect beneficiaries' privacy, the LDS does not include the zip code of residence for beneficiaries. Instead, the county of residence is provided. HRRs are defined based on zip codes of residence, so the LDS does not contain enough information to

assign an HRR of residence for each beneficiary.

To get around this problem, a mapping of counties to HRRs was created. Both HRRs and counties can be defined by zip code relatively easily. Zip codes are nested in HRRs. Zip codes are nearly nested in counties.¹ Unfortunately, HRRs and counties are not nested. A single HRR can contain multiple counties, and conversely, a single county can contain multiple HRRs. To create a county-to-HRR crosswalk, census data on the number of people aged 65 and older living in each zip code was summed for each county. The proportion of each county's population living in each HRR was then tabulated. If at least 70 percent of a county's population lived in a particular HRR, that HRR was assigned to the entire county.

In some cases, counties were fairly evenly distributed across multiple HRRs. Some HRRs were grouped together to facilitate assignment of a county to a single HRR group. For example, the Chicago area has several HRRs in a small area and some very large counties. As a result, the Cook County Medicare-age population is spread across 7 HRRs: 44% in the Chicago region, 20% in the Evanston region, 18% in the Blue Island region, 14% in the Melrose Park region, 2% in the Elgin region, 2% in the Munster, IN region, and less than 1% in the Joliet region. HRRs were merged in the Chicago area and in other areas with similar issues on a case-by-case basis until all counties could be assigned to a single HRR group. Counties with very small populations (less than 7500 people aged 65 and older) were generally assigned based on the HRR of the majority of the population, without merging HRRs. A full list of the modified HRRs is available in Appendix D. These modified HRRs are referred to in this paper as HRR groups. About one-quarter of HRRs were aggregated: the 308 HRRs defined by the Dartmouth Atlas are grouped into 226 HRR groups for this analysis. Figure 3.2 illustrates the grouping of HRRs. White HRRs are unchanged

1. There are a small number of zip codes that fall within multiple counties. In these cases, I selected a single county for each zip code based on where the majority of the zip code's population lives.



White HRRs retain their original Dartmouth Atlas borders. Shaded HRRs are grouped. HRR map boundary layer file from Esri (2012) using Dartmouth data.

Figure 3.2: Dartmouth Atlas hospital referral regions: grouped for use with county-level data.

from their original Dartmouth Atlas definition. HRRs in shades of grey are grouped. The different shades are for clarity where multiple grouped HRRs share borders; there is no meaning in the particular shade of grey.

The final county-to-HRR crosswalk misclassifies a very small proportion of the total US age 65-plus population; 2.2% of people were assigned to a different HRR group using county of residence than would be assigned based on zip code of residence (using Census data). These misclassified people tend to be located in areas with multiple HRRs nearby—these are the areas where the original Dartmouth Atlas classification would be expected to have a lot of error, as well, so the new classification may not introduce much more error.

Measures

ED visits were identified using the Carrier and Institutional files of Medicare data. Only visits that included an institution charge were included, following the definition that seems to have been used in the Medicare Geographic Variation PUF (CMS 2015). An additional 162,138 ED visits are identifiable in the Medicare administrative data for the aged FFS cohort that have only a physician charge from an ED, but no associated institutional claim in the Inpatient or Outpatient files. These visits were dropped from the analysis.

Primary care visits were defined using the CMS list of primary care specialties and primary care Evaluation and Management codes (CMS 2012d); primary care specialties include general practice, family medicine, internal medicine, geriatric medicine, pediatrics (not applicable to this population), nurse practitioners, physician assistants, and clinical nurse specialists. A physician's specialty is taken from the Carrier claim. A visit is defined as one or more primary care E&M CPT codes charged by the same physician on a single day; see Appendix C for a list of E&M codes used to identify primary care visits. A maximum of one visit is counted per person per day; that is, if a beneficiary was seen for two separate primary care E&M visits in a single day, only one is counted.

In the individual-level analyses, "treatment" is a binary variable indicating whether a beneficiary had a primary care relationship. This is operationalized as a beneficiary having at least two primary care visits with a single primary-care physician, as defined by the NPI number over the time we observe him/her. Most beneficiaries—76%—were included in the data for two full years and the average tenure of beneficiaries in the sample is 22.1 months, so this generally amounts to requiring an average of one primary care visit per year with the same physician.

HRR-level analysis

Models in the HRR-level analysis included HRR group-level covariates to adjust results for variation across the HRR groups in primary care physician supply, beneficiary age, race, sex, Medicaid eligibility (percent of HRR group beneficiaries), comorbidities using the average HCC score, and area intensity of practice using the Dartmouth Atlas end of life care intensity index. HRR group level variables were created based on data from the Medicare Geographic Variation Public Use File (CMS 2015), the the Dartmouth Atlas (Dartmouth Atlas 2015), and the Area Health Resource File (AHRF; HRSA (2014)). When HRRs were aggregated into HRR groups, the HRR level variables were summarized with a weighted average of the values for each HRR in the group (weighted by the age 65 and up population). Table 3.1 presents summary statistics for the HRR groups in the analysis.

Table 3.1: Summary statistics (2011): HRR groups

	mean	sd	min	max
Aged FFS Medicare beneficiary population	5991.8	8226.8	694	71241
Age (average)*	75.7	0.74	74	77.3
Female (percent)*	56.6	1.58	51.1	60.3
Non-Hispanic White (percent)*	86.7	11.8	26.0	98.6
Intensity of care index**	0.86	0.23	0.46	1.84
Mortality rate*	4.99	0.50	3.50	6.37
Comorbidity index (average HCC score)*	0.95	0.071	0.75	1.24
Primary care phys per 100k population***	124.7	29.1	66.4	262.2
Primary care visit rate (annual visits per beneficiary)	3.85	0.64	2.20	5.71
ED visits per 100 benes	55.9	7.18	36.7	77.5
PCT ED visits per 100 benes	9.50	3.32	3.34	27.9
Ballard non-emergent ED visits per 100 benes	14.0	2.64	8.47	25.8
<i>N</i>	226			

Sources: * from Medicare Geographic Variation PUF; ** from Dartmouth Atlas; *** from AHRF

Counts of ED visits and primary care visits were taken from the 2011-12 Medicare

data and aggregated at the HRR group and year level, including the number of ED visits, the number of ED visits rated as primary care treatable using the E&M algorithm, the number of ED visits rated non-emergent by the Ballard version of the Billings/NYU algorithm, and the per capita annual rate of primary care visits by the cohort of Medicare beneficiaries.

Generalized estimating equations (using the poisson family and a log link) were used to estimate the effect of the average number of primary care visits per beneficiary per year on the number of ED visits. The number aged FFS beneficiaries living in each HRR group was included in the equation as an exposure variable, with the coefficient on the log-population constrained to 1. Thus the exponentiated coefficients from the model are interpreted as the ratio of incidence rates of ED visits per person per year. The population-averaged estimator was specified, with an exchangeable working correlation matrix and the Huber/White variance estimator.

Individual-level analysis

As noted above, the individual-level analyses include a binary treatment variable indicating whether a beneficiary had a primary care relationship (at least two primary care visits with a single primary-care physician over the time we observe him/her). The strength of the primary care relationship with the most frequently seen physician is measured as the proportion of all of a beneficiary's primary care visits that are with the modal physician. The number of days with a primary care visit, potentially an indicator of how sick a person is or possibly how much he/she enjoys visiting the doctor, is measured in quintiles. Summary statistics at the beneficiary level are in Table 3.2.

The outcome of interest in most models is the number of primary care treatable ED visits by a beneficiary in a particular year, as identified using the new algorithm for identifying primary care treatable ED visits. In the treatment bounds model, the

Table 3.2: Summary statistics: Individual beneficiaries

	mean	stdev	min	max
Age	74.4	8.3	63	98
% visits with modal PC physician	68.5	36.5	0	100
HCC community score	1.18	1.24	0.29	18.28
PC trtable ED visits/yr	0.087	0.371	0	109
All ED visits/yr	0.629	1.35	0	126
Primary care visit days/yr	4.46	5.14	0	275
Had no PCT ED visits (bene-year)	93.11%			
Female	57.01%			
White	86.39%			
Medicaid	14.02%			
Had primary care relationship	73.58%			
HRR group level variables				
Intensity index	1.01	0.281	0.46	1.84
Prim Care Phys per 100k	127.5	25.8	66.4	281.1
Total Phys per 100k	254.6	72.81	105.3	784.1
N = 1,475,929 beneficiaries				

outcome of interest is whether the beneficiary had any primary care treatable ED visits in the year. In the model checks, the outcome of interest is the number of weekend ED visits in a year or the number of injury-related ED visits in a year.

Other included variables include age, gender, race (white vs. non-white), comorbidities as measured using the HCC community score (`score_community`), the end of life care intensity index from the Dartmouth Atlas for the HRR group of reported residence, total physician supply and primary care physician supply measured as physicians per 100,000 population for the HRR group of reported residence. County-level primary care physician counts were taken from the 2011 and 2012 AHRF and defined as the number of non-federal primary physicians engaged in patient care and excluding hospital residents minus the pediatricians, plus physician assistants, nurse practitioners, and clinical nurse specialists, then summed by HRR group (HRSA

2014).

All analysis was performed with Stata version 14 (StataCorp 2015a).

Results

HRR-level analysis

Regression analysis shows a modest reduction in primary care treatable ED visits associated with an increase in the average number of primary care visits per Medicare beneficiary per year. Table 3.3 presents the results of negative binomial regressions of the effect of primary care visit rates on ED visit outcomes. Column 1 displays results for primary care treatable ED visits; the incidence rate ratio of 0.852 (95% CI: 0.805 to 0.901) suggests that one additional primary care visit per year is associated with a 14.8% decrease in the rate of primary care treatable ED visits per beneficiary. The estimated incidence rate of primary care treatable ED visits (not shown in table) is 12.1 visits per 100 beneficiary-months at the lowest observed level of primary care visits: 2 primary care visits per beneficiary (95% CI: 10.9 to 13.4 PCT ED visits per 100 beneficiary-months). At the highest level of observed primary care visit rates, 6 primary care visits per beneficiary per year corresponds with a predicted incidence rate of primary care treatable ED visits of 6.4 per 100 beneficiary-months (95% CI: 5.5 to 7.3 PCT ED visits per 100 beneficiary-months), about a 47 percent reduction. By contrast, results suggest that increased primary care visit rates have a small effect on the total number of ED visits by aged FFS Medicare beneficiaries (IRR 0.974, 95% CI: 0.954 to 0.995; Column 2 of Table 3.3), and non-emergent ED visits from the Billings algorithm show a small effect (IRR 0.914, 95% CI: 0.885 to 0.945; Column 3 of Table 3.3).

Because the units of observation are geographies, a spatial analysis could be more

Table 3.3: HRR-level analysis results

	(1) Primary care trtble ED visits	(2) Total ED visits	(3) NYU: non-emergent ED visits
Primary care visit rate	0.852*** [0.805,0.901]	0.974* [0.954,0.995]	0.914*** [0.885,0.945]
Mortality rate	1.175*** [1.070,1.289]	1.128*** [1.086,1.172]	1.156*** [1.099,1.217]
Avg age	1.088** [1.031,1.148]	0.977* [0.958,0.997]	0.999 [0.972,1.027]
Female (percent)	1.034** [1.011,1.057]	1.011* [1.001,1.021]	1.012 [0.998,1.027]
Non-Hispanic White (percent)	1.001 [0.997,1.005]	1.000 [0.998,1.002]	0.998* [0.996,1.000]
Intensity of care index	0.878 [0.712,1.083]	0.903** [0.839,0.972]	0.877 [0.767,1.003]
Primary care phys per 100k population	1.002*** [1.001,1.003]	1.001*** [1.001,1.002]	1.002*** [1.002,1.003]
Medicaid eligible (percent)	1.015*** [1.008,1.023]	0.999 [0.996,1.002]	1.003 [0.999,1.007]
Comorbidity index (avg HCC score)	0.323 [0.102,1.020]	2.744*** [1.919,3.922]	1.661 [0.910,3.033]
<i>N</i>	452	452	452

Exponentiated coefficients; 95% confidence intervals in brackets

All variables defined as percents are multiplied by 100 to give IRRs for a 1 percentage point change

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

appropriate, particularly if there are strong regional effects on primary care visit rates or primary care treatable ED visit rates. A separate analysis using a spatial error model did not yield substantially different results. See Appendix E.

Person/year-level analysis

To facilitate interpretation, the results of individual analyses are presented in Table 3.4 as marginal effects, rather than coefficients from the different types of analyses, which cannot be readily compared. The marginal effects are the reduction or increase in the mean count of the outcome per 100 person-years for people with a primary care relationship compared to people without a primary care relationship—i.e., the treatment effect measured in units of ED visits. A negative treatment effect indicates that people with a primary care relationship had fewer ED visits. A percent decrease is also given for context; it is the treatment effect divided by the predicted count for people without a primary care relationship. They are shown separately for each analysis. Results from the bounded treatment effects analysis are presented separately because of the difference in the outcome variables between that analysis and all the others.

All the models in Table 3.4 showed a negative effect. The “naïve” model estimated a reduction of 3.62 primary care treatable ED visits per 100 beneficiary-years (a 30.5% reduction), while the endogenous treatment poisson estimated a reduction of 3.16 primary care treatable ED visits (a 27.5% reduction). The structural equation model version of the endogenous treatment estimate found a similar effect of 2.72 PCT ED visits per 100 beneficiary-years, which represented an 8.12% reduction.

The zero-inflated negative binomial model attempting to control for selection effects by looking only at people who had demonstrated a willingness to use the ED in the previous year found a similar effect to the endogenous treatment poisson: a reduction of 7.31 primary care treatable ED visits per 100 beneficiary-years repre-

senting a 30.9% decrease. (The larger absolute figure of 7.31 visits versus 3 to 4 visits in the models looking at the full set of beneficiaries reflects a larger baseline number of PCT ED visits in the population with any ED visits in the previous year.)

The model testing for an availability effect by looking at the effect of a primary care relationship on primary care treatable ED visits on the weekend found an effect fairly similar (in the reduction proportion) to the endogenous treatment poisson: 21.7% in weekend visits (and 27.5% in PCT visits). The marginal effect in units of visits was, of course, smaller for weekend visits, since the baseline number of weekend PCT visits per capita is much smaller than the total number of PCT visits.

The estimation check looking at the effect of a primary care relationship on visits that should be less sensitive to that relationship found a moderate increase in ED visits for accidents associated with having a primary care relationship, potentially. This could reflect a problem with our definition of accidents or could be due to other lingering selection issues. It's possible that people who have primary care relationships are more likely to seek physician care for minor accidents.

The bounded treatment effects analysis is useful to put these other results in context. See Table 3.5. The columns are the two assumptions about error in determining treatment status. Column 1 has the range of results assuming that both false negatives (a person incorrectly classified as having no primary care relationship) and false positives (a person incorrectly classified as having a primary care relationship) are possible. Column 2 assumes that only false negatives are possible. Panel A contains the results for the exogenous selection assumption for a range of treatment status error rates from 0 to 10 percent. The bracketed ranges are not confidence intervals, they are point estimates (and labeled p.e. to remind the reader, since they look like CIs) for bounds on the treatment effect under each set of assumptions. Recall that the outcome in this analysis is binary and has a positive connotation: it is 1 if the beneficiary had no primary care treatable ED visits during the time he or she was

observed and 0 otherwise. Column 2 of Panel A gives the bounds of the causal treatment effect assuming no false positives in determining treatment status and assuming an exogenous selection process. Row 1 of that Panel indicates that, assuming no error in determining treatment status, the causal treatment effect is -0.022 (note, the lower and upper bound are the same, as this assumes no error in treatment status and no selection effects); this means that (under these assumptions) having a primary care relationship causes a 2.2 percentage point reduction in the probability that a Medicare beneficiary has no primary care treatable ED visits. In other words, having a primary care relationship makes the beneficiary more likely to have a primary care treatable ED visit. To put the magnitude of the effect in context, 93.11% of beneficiary-years had no PCT ED visits, so 2.2 percentage points is a very small treatment effect.

Note that as the treatment status error rate goes up from 0 to 10 percent, the range of possible causal treatment effects gets wider. With 10 percent error in treatment status (and assuming exogenous selection and no false positive treatment status errors), the causal treatment effect of having a primary care relationship must lie between -9 and 2.3 percentage points.

This table tells us that we can only identify the sign of the treatment effect under some of the assumption scenarios. Cells in italics identify a detrimental treatment effect. Cells in bold face identify a beneficial treatment effect. Plain cells cannot tell us whether the effect is beneficial or detrimental. Assuming we have positive selection—that is, people who choose to have a primary care relationship are less likely to have primary care treatable ED visits—makes it nearly certain that having a primary care physician causes people to have more primary care treatable ED visits. In contrast, if we assume no treatment status error, negative monotone treatment selection (MTS), and that beneficiary age is a suitable monotone IV, we can be certain that the treatment effect is positive and lies between 0.7 and 58.7 percentage points.

If we assume that there is an endogenous selection process, Panel B tells us that, depending on the error rates and direction in determining treatment status, the causal treatment effect must lie between -38.8 percentage points and positive 81.2 percentage points. We can choose the assumptions we find most likely to further narrow down the possible range of treatment effects. Assuming negative monotone treatment selection and a low treatment status error rate of 0 to 2 percent, we can use Panel E to state that the treatment effects lies between a small negative effect of 8.4 percentage points and a large positive effect of 62.7 percentage points.

Table 3.4: Person/year-level analysis: marginal effects (reduction/increase in mean count of outcome per 100 person-years)

Model	Population	Outcome	Marginal effect (visits per 100 bene-yrs)	95% CI	per- cent de- crease
Naïve zero-inflated neg. binomial	Aged FFS	Prim care trtable ED visits	-3.62	[-4.12,-3.13]	-30.5%
Endog. trtmt Poisson	Aged FFS	Prim care trtable ED visits	-3.16	[-4.22,-2.10]	-27.5%
SEM Endog. trtmt	Aged FFS	Prim care trtable ED visits	-2.72	[-3.86,-1.59]	-8.12%
Zero-inflated neg. binomial	Aged FFS with prev. yr. ED visit	Prim care trtable ED visits	-7.31	[-9.56,-5.06]	-30.9%
Endog. trtmt Poisson	Aged FFS	Prim care trtable weekend ED visits	-0.36	[-0.56,-0.17]	-21.7%
Endog. trtmt Poisson	Aged FFS	ED visits for accidents	0.33	[0.07,0.59]	13.4%

All models adjusted for age, race, sex, Medicaid status, HCC community score (as quintiles), total PC visits (as quintiles), and proportion of PC visits with modal PC physician.

Zero-inflated models have the same control variables in the main model and the inflation model. Endogenous treatment models have the same control variables in the main model and the treatment model except that treatment model additionally includes three HCC group-level variables: Dartmouth treatment intensity index, AHRF PC physicians per 100k population, and AHRF total doctors per 100k population.

Table 3.5: Bounded treatment effect analysis results

Outcome: No primary care treatable ED visits during the time observed		
	(1)	(2)
Error Rate	Arbitrary Errors	No False Positives
A. Exogenous Selection Model		
0	<i>[-0.022, -0.022]</i> p.e.	<i>[-0.022, -0.022]</i> p.e.
0.01	[-0.074, 0.028] p.e.	<i>[-0.074, -0.019]</i> p.e.
0.02	[-0.091, 0.075] p.e.	<i>[-0.090, -0.015]</i> p.e.
0.05	[-0.095, 0.205] p.e.	<i>[-0.090, -0.004]</i> p.e.
0.1	[-0.101, 0.269] p.e.	[-0.090, 0.023] p.e.
B. No Monotonicity Assumptions (Worst Case Selection)		
0	[-0.288, 0.712] p.e.	[-0.288, 0.712] p.e.
0.01	[-0.298, 0.722] p.e.	[-0.298, 0.722] p.e.
0.02	[-0.308, 0.732] p.e.	[-0.301, 0.732] p.e.
0.05	[-0.338, 0.762] p.e.	[-0.301, 0.762] p.e.
0.1	[-0.388, 0.812] p.e.	[-0.301, 0.812] p.e.
C. MTS Assumption: Negative Selection		
0	[-0.022, 0.712] p.e.	[-0.022, 0.712] p.e.
0.01	[-0.074, 0.722] p.e.	[-0.074, 0.722] p.e.
0.02	[-0.091, 0.732] p.e.	[-0.090, 0.732] p.e.
0.05	[-0.095, 0.762] p.e.	[-0.090, 0.762] p.e.
0.1	[-0.101, 0.812] p.e.	[-0.090, 0.812] p.e.
D. MTS Assumption: Positive Selection		
0	<i>[-0.288, -0.022]</i> p.e.	<i>[-0.288, -0.022]</i> p.e.
0.01	[-0.298, 0.028] p.e.	<i>[-0.298, -0.019]</i> p.e.
0.02	[-0.308, 0.075] p.e.	<i>[-0.301, -0.015]</i> p.e.
0.05	[-0.338, 0.205] p.e.	<i>[-0.301, -0.004]</i> p.e.
0.1	[-0.388, 0.269] p.e.	[-0.301, 0.023] p.e.
E. MIV and MTS Assumptions: Negative Selection		
0	[0.007, 0.587] p.e.	[0.007, 0.587] p.e.
0.01	[-0.060, 0.607] p.e.	[-0.060, 0.597] p.e.
0.02	[-0.084, 0.627] p.e.	[-0.083, 0.607] p.e.
0.05	[-0.088, 0.687] p.e.	[-0.083, 0.637] p.e.
0.1	[-0.095, 0.760] p.e.	[-0.083, 0.687] p.e.
F. MIV and MTS Assumptions: Positive Selection		
0	<i>[-0.287, -0.049]</i> p.e.	<i>[-0.287, -0.049]</i> p.e.
0.01	<i>[-0.298, -0.011]</i> p.e.	<i>[-0.297, -0.047]</i> p.e.
0.02	[-0.308, 0.025] p.e.	<i>[-0.301, -0.044]</i> p.e.
0.05	[-0.338, 0.117] p.e.	<i>[-0.301, -0.037]</i> p.e.
0.1	[-0.388, 0.168] p.e.	<i>[-0.301, -0.024]</i> p.e.

p.e.=point estimates; MTS=Monotone treatment selection; MIV=Monotone instrumental variable

Bold text indicates an assumption scenario identifies a positive treatment effect (i.e., people with primary care relationship less likely to have a PCT ED visit); *italics* indicate an assumption scenario identifies a negative treatment effect

Discussion

The expected association between primary care use and decreased rates of primary-care treatable ED visits was observed in this data both at a hospital referral region level and at an individual level using several different methods. However, a bounded treatment effects analysis was unable to definitively determine the sign on the causal treatment effect, leaving it an open question whether our analyses succeeded in identifying a true treatment effect or whether there was residual bias.

At the HRR level, we estimated a substantial reduction of nearly 15 percent in rates of primary care treatable ED visits associated with an increase in the number of primary care visits per beneficiary per year. The reduction in visits was concentrated in primary care treatable visits; the overall reduction in ED visits was much smaller at 2.6 percent.

When the Billings/NYU algorithm was used to identify non-emergent ED visits, a smaller effect was found between higher rates of primary care visits and rates of less severe ED visits: an 8.6 percent reduction. This may suggest that the Minnesota algorithm is more sensitive for identifying the kind of less-severe ED visits that are susceptible to increased access to primary care. Further analysis is needed to understand the differences between the algorithms.

The individual-level analysis found a large reduction in primary care treatable ED visits associated with having a primary care relationship. However, this analysis is subject to significant selection effects that may not have been completely adjusted for by our endogenous treatment models. The variable included in the treatment equations as plausibly exogenous to the outcome equation except through the treatment variable does not have ideal characteristics. An ideal instrument has a large effect on the treatment variable in the expected direction. Our candidate instrument the supply of primary care physicians has a small effect on the probability of a primary

care relationship and the effect goes in the wrong direction. That is, an increased supply of primary care physicians is associated with a decreased probability that an individual has a primary care relationship. This makes our results somewhat suspect.

We found a moderate and unexpected increase in ED visits caused by accidents associated with primary care. We had expected to find little or no impact of primary care relationships on ED visits for accidents, which would serve as a test of our identification strategy. Instead, we found an increase in these visits of 13.4%. This could suggest that our identification strategy failed. However, it could represent a true, but unexpected finding. We used E-codes to identify ED visits for accidents. However, E-codes do not specify the severity of an accident. It could be that people with primary care relationships seek care for minor accidents that others treat at home.

We also found a large reduction in the expected number of weekend primary care treatable ED visits associated with having a primary care relationship. This could mean that availability of primary care is not a major mechanism for the reduction in PCT ED visits associated with having a primary care relationship. However, it could instead mean that people with a primary care relationship are more likely to wait until they can see their regular physician rather than going to the ED on a weekend, when their problem allows them to wait.

It is important to consider whether all of the effects found in our analyses reflect a residual difference in health between people with and without a primary care relationship. That is, people with a primary care relationship are likely to be sicker than people without a PC relationship, all else being equal, and sicker people are less likely to have primary care treatable ED visits. We can't definitively reject this hypothesis. However, it is reassuring that unadjusted rates of primary care treatable ED visits are, in fact, higher in people with primary care relationships.

Limitations

The estimation and interpretation of the individual-level analyses are hampered by the lack of a good instrument for having a primary care relationship: that is, a variable that is associated with having a primary care relationship but not with unmeasured confounders in the relationship between primary care relationships and primary care-treatable ED visits. A candidate instrument the supply of primary care physicians in each geographic area displays the opposite relationships expected. The model specifications intended as falsification tests did not have the expected results, and the bounded treatment effect analysis did not determine the sign of the treatment effect, leaving us uncertain of the success of our identification strategy.

Conclusions

This study presents evidence from observational data that receipt of primary care is associated with a reduction in primary care treatable ED visits, though residual selection effects cannot be ruled out. It is the first study to use the new Minnesota algorithm to identify these less severe visits. Both hospital referral region-level and individual-level analyses suggest a positive impact of increased use of primary care on primary care treatable ED visits. In economics parlance, the individual-level analysis suggests a positive impact on the extensive margin of primary care use, while the geographic analysis suggests a positive impact on the intensive margin. This study adds to the evidence from small-scale randomized trials of primary care interventions, suggesting that this beneficial effect is detectable in observational data at a population level.

Chapter 4

Paper Three: Healthcare service use among elderly seasonal migrators

A non-trivial proportion of American elders migrate seasonally, traveling between a summer and winter home. Little is known about their health care seeking behaviors. This topic is timely, as ACOs that serve Medicare beneficiaries will be responsible for their care wherever they are. Using Medicare data, we identify seasonal migrators and compare the care they receive in their summer homes to that received in their winter homes, as well as to care received by beneficiaries who do not migrate. Previous studies have used this technique to identify seasonal migration, but have not reported on primary care use by location in the Medicare population (Buczko 1992, 1994; Al-Haque et al. 2015).

Having a primary care relationship is associated with a modestly reduced use of the emergency department for primary care treatable complaints (see paper two of this dissertation). In this analysis, we explore the effect of having local primary care on the use of other healthcare services. For this purpose, local primary care is defined as seeing a nearby primary care physician for an outpatient evaluation and management visit. “Nearby” is generally defined as within 250 miles of the county where a beneficiary resides; see below for full details. We use location information

from claims data to determine where each beneficiary lives and where he or she receives primary care and other health services.

We have not found any previous studies that have reported what proportion of seasonal migrators have a primary care relationship in each location where they live. We hypothesized that Snowbirds (who winter in warm areas) would consider their northern home to be their primary home, and thus would be more likely to have a primary care relationship in the north. We further hypothesized that they may be more likely to experience a primary care treatable ED visit in their winter home, due to their lack of a primary care relationship there.

Methods

Description of sample

The data source was the Medicare Limited Data Set (LDS) five percent sample for the years 2011 and 2012. Aged fee-for-service beneficiaries who stated their residence to be in the fifty United States plus Washington, D.C., were included. Beneficiaries were required to have no months of HMO coverage and equal months of hospital and supplementary medical insurance to ensure that claim data was as complete as possible a picture of total health care use. The 24 months of claims data were summarized into eight seasons: January to March (winter), April to June (spring), July to September (summer), and October to December (fall) for 2011 and 2012. There were 1,483,567 beneficiaries eligible for the analysis. Most received some medical care over the course of the time they were observed; 5.19% received no location-attributable care. In total, we observed the location of a beneficiary (through medical care or the denominator file) in 81.67% of the seasons in which they were eligible for the analysis.

Location information in claims data

Medicare data gives only one official place of residence for each beneficiary per year: their stated location as of December of that year. Other sources of location information for each beneficiary came from claims data. Each claim includes the county where the beneficiary said he/she lived at the time the claim was submitted. Organization (e.g., hospital) and physician National Provider Identifier (NPI) numbers were linked to the National Plan and Provider Enumeration System (NPPES) file provided by the National Bureau of Economic Research (NBER) to determine the zip code where care was given (Roth 2014).

In some rare cases, physicians were observed to be giving care hundreds of miles from their stated office location. A spot check of these cases suggested that these physicians had moved since the NPPES file was published. We used other claim information where possible to identify physicians who had moved and did not use their post-move care to determine beneficiary locations. For example, if the NPPES file listed a physician as practicing in Hennepin County, Minnesota, but more than half of his or her patients reported living more than 500 miles from Hennepin County, we assumed the physician had moved and didn't use the Hennepin County practice address to determine beneficiary locations for that year.

Several indicators of care were summarized by season and county including: the number of claim line-items with that location listed as the beneficiary residence, the number of inpatient days, outpatient days, hospice service days, SNF days, days receiving physician services, days receiving an outpatient E&M service from a primary care physician, and days receiving an outpatient E&M service from a specialist. Primary care specialties were defined to be general practice, family medicine, internal medicine, geriatrics, nurse practitioner, Clinical Nurse Specialist (CNS), and physician assistant. All other physicians were considered specialists. This definition is taken from a CMMI primary care initiative (CMS 2012d).

The distance from the beneficiaries' stated county of residence (from the denominator file) to the county where care was provided was calculated using data from NBER (Roth 2015). For each season in which each beneficiary was eligible for the analysis, the care indicators were summarized according to how far the county of care was from the county of residence. Each beneficiary was assigned a location for each day on which care was received. If the beneficiary lived within 250 miles of the location of care (measured by distance between counties), he or she was considered to be at home. If claims information placed the beneficiary in multiple counties on the same day and any of those counties was within 250 miles of the beneficiary's home, he or she was considered to be at home. In other words, if any evidence suggested a beneficiary was within 250 miles of home, we threw out information that seemed to place the beneficiary farther from home.

Claim data provided a relatively complete picture of beneficiaries' locations throughout the year for most beneficiaries in the sample. Of the total 8.7 million person-seasons of eligibility in the two-year sample, 18.6% provided no information on the beneficiary's location, either because that beneficiary received no care or because the location of the care was undetermined due to incomplete data in the NPI file. The other 81.4% of person-seasons observed had at least one piece of information from claims to help determine where the beneficiary was located during that season.

People who had no claims information for an entire year were included in the analysis with their denominator file residence as a fall (October–December) observation because the denominator file residence is as of December.

Beneficiaries were allowed to be in multiple geographies in a single season. Each beneficiary–season–geography observation was included in the analysis. Seasons in which beneficiaries received no care were not included in the analysis, except as noted for people with no care through the entire year.

Some beneficiaries were poorly tracked by claims data because they rarely (or

never) received health care. This group could include people who are very healthy, people who receive care through the VA system, and people who receive care that is not covered by Medicare (for example, some types of complementary medicine like acupuncture). We categorized these people using the information available to us; their migration category is likely less accurate than beneficiaries for whom we had more complete data.

Carrier file data for providers other than physicians was not included in the analysis used to determine beneficiaries' locations. This was done because laboratories and ambulance providers billing on the carrier file could be located far from where the beneficiary was actually receiving care.

Other discrepancy adjustments: There were approximately 112,000 beneficiary–days (out of more than 61 million) where claims for medical care received on a particular day had different beneficiary counties of residence. In those cases, the county located closest to the physician's location was kept and the others discarded. This seemed to happen more frequently with beneficiaries who have multiple residences (e.g., Snowbirds).

We have a single year of NPI location information (2012); we assume that hospitals and other institutional providers locations are more stable in location than individual physicians, who are more likely to have moved and thus have inaccurate location data in the NPI file. For this reason, location information from the institutions (inpatient, outpatient, SNF) was used in preference to other types of information for each day that the patient was in residence or treated at the institution.

Classification of beneficiaries' seasonal migration patterns

Locations of care and residence were classified by a state/region scheme. The Bureau of Economic Analysis (BEA) regions were used with a few modifications (BEA 2015). The Far West BEA region was divided into northern (AK, WA, OR) and southern

(CA, HI, NV) Far West regions. States that are of particular interest as donors or recipients of Snowbirds were considered separately from the rest of their BEA region. This state/region scheme yields 23 geographic entities that were compared in the analyses (referred to as “locations”). States and regions considered to be in the north for purposes of migration classification were BEA regions New England, Mideast, Great Lakes, Plains, Rocky Mountain, and the northern part of the Far West region, and Connecticut, Massachusetts, Michigan, Minnesota, New Jersey, New York, Pennsylvania, and Wisconsin. All other states and regions were considered to be in the south.

People who received all their medical care in northern geographies were classified as North only; people receiving all care in southern geographies were classified as South only (46.8%). People who received some winter or spring care in the south and some fall or summer care in the north were classified as Snowbirds (4.1%). People who didn't fit any of the other patterns were classified as “Other” (1.4%); the “Other” category also includes people who had no medical care, but changed their address from a northern location to a southern location (or vice-versa) from 2011 to 2012.

We don't have a strong sense of the people in the “Other” category. They are reverse Snowbirds, in essence, spending the hot season in the south and the cold seasons in the north. They may not be true migrators; instead, they may have (permanently) moved from the south to the north or vice versa or they may have received healthcare services while on vacation or on a short visit to another part of the country.

Measures

Six measures of healthcare use are explored: primary care treatable ED visits (as categorized by the Minnesota algorithm; see Paper One of this dissertation), total ED visits, days of inpatient hospitalization, visits with specialist physicians (includes

a maximum of one per day), days in a skilled nursing facility (SNF), and days of hospice care. Each measure is tallied by beneficiary, location, and season.

Modeling

Population-averaged negative binomial regression was used to model the counts of healthcare services used by season and location for each eligible beneficiary. Beneficiary-seasons with no observed location information were not included in the analyses. That is, we did not impute any beneficiary locations during months in which they did not receive care (with the exception described above for people receiving no care during an entire year).

We included the log of the number of months per season with any healthcare in the regression as an exposure variable (that is, with the coefficient constrained to one). This provides some normalization, so that the incident rate ratios estimated by the regressions can be interpreted as the number of incidents per month of care received.

The model was estimated using generalized estimating equations, using the `xtbreg` command in Stata 14 (StataCorp 2015a). The Huber–White “sandwich” variance estimator was specified, as was an exchangeable working correlation matrix.

All models included main effects for migration category, having primary care at the location of the observation, location (region or state, as described above), beneficiary age, race (white vs. non-white), sex, Medicaid status (any months of state buy-in coverage), season, and a measure of comorbidities (quintiles of the Hierarchical Condition Category community score). Interaction terms crossing migration category by primary care status are also included. Results are presented as incidence rate ratios per month receiving any healthcare. Predicted rates of PCT ED visits and total ED visits per month of healthcare receipt are also provided by migration category–primary care status groups to facilitate interpretation.

Results

Table 4.1 shows the characteristics of the various migration groups. Overall, Snowbirds make up 4.11% of the population. The Other migration category is much smaller: 1.35% of the population. People living in the North only and in the South only make up relatively equal proportions of the remainder of the population.

Proportions and means in Table 4.1 that are not significantly different at the $p=0.01$ level are indicated. Large sample sizes result in some statistically significant difference that have little practical significance (for example, the proportion who are female in the North only and South only groups). Some very clear differences are apparent. Snowbirds were more likely to be white and less likely to be on Medicaid than other migration categories. By design, all snowbirds had to have received some medical care, but they also have a larger proportion of seasons during which they received any medical care.

Furthermore, both Snowbirds and people in the Other category are more likely to have received any primary care. This difference persists when the sample is limited to beneficiaries with a full 24 months of eligibility for the analysis and when the sample is limited to beneficiaries who received any medical care over the course of the study (figures not shown).

Table 4.1: Sample characteristics

	Migration Categories				
	Total sample	North only	South only	Snow-birds	Other
Number in sample	1,483,567	708,203	694,299	61,008	20,057
Percent of population ¹	100%	47.74%	46.80%	4.11%	1.35%
Female	56.99%	57.44% ^A	56.64%	55.57%	57.45% ^A
White	86.31%	89.09%	82.76%	93.85%	88.25%
Medicaid ²	13.21%	11.32%	15.90%	5.50%	9.95%
Beneficiaries receiving no medical care ³	5.19%	5.13%	5.82%	–	1.12%
Person–seasons with any medical care ⁴	81.67%	81.35%	80.99%	91.09%	85.09%
Eligible months ⁵ (sd)	21.08 (5.98)	21.03 (6.04)	20.93 (6.10)	23.03 (3.12)	21.76 (5.21)
Age (sd)	74.10 (8.30)	74.51 (8.51)	73.72 (8.17)	74.13 (7.26)	73.12 (7.64)
Comorbidities: average HCC community score (sd)	1.19 (1.17)	1.19 (1.18)	1.18 (1.17)	1.23 (1.10)	1.26 (1.24)
Has any primary care	82.2%	81.4% ^A	81.5% ^A	95.4%	91.7%
Has primary care in N	42.8%			74.4%	63.1%
Has primary care in S	41.9%			69.6%	62.0%
Has primary care in both N and S	2.5%			48.7%	33.4%
Observations: n=1,483,567 unique beneficiaries, N=8,892,882 bene/season/locations					

^A Estimates not significantly different from each other at p=0.01 (within row)

¹ Percentage of aged FFS Medicare beneficiaries

² Beneficiaries receiving any months of state buy-in coverage

³ Percentage of beneficiaries receiving no location-attributable medical care during the entire time they were eligible for the analysis; note that the definition of migration categories precludes any Snowbirds with no medical care

⁴ Percentage of total eligible beneficiary-seasons in which any location-attributable medical care was received

⁵ Months of parts A and B coverage received as recorded in denominator file

Table 4.2 gives unadjusted rates of healthcare service use per month of eligibility for inclusion in the analysis. These are the rates as observed, not adjusted for beneficiary characteristics (see Table 4.3 for adjusted estimates). Given the large sample, rates that are not different at $p=0.01$ are indicated.

The biggest difference in healthcare use between the migration categories was in days of hospice care. The migratory groups (Snowbirds and Other categories) had much lower rates of hospice use than the non-migratory groups (North only and South only).

Both migratory groups had higher rates of ED visits and primary care treatable ED visits than the non-migratory groups. They also had more visits with specialist physicians, more days with outpatient services, and more days seeing a physician in any setting.

Interestingly, despite their having higher rates of any primary care visit, Snowbirds had the lowest number of primary care physician visits (26% lower than the overall average).

Table 4.2: Unadjusted healthcare service use rates (not adjusted for beneficiary characteristics)

	Migration Categories				
	Total sample	North only	South only	Snowbirds	Other
Average days or visits per month eligible for analysis [95% C.I.] ¹					
Inpatient days	0.184 [0.183,0.185]	0.182 [0.181,0.183]	0.184 ^A [0.183,0.186]	0.188 ^A [0.184,0.192]	0.209 [0.201,0.216]
Hospice days	0.193 [0.190,0.195]	0.178 [0.174,0.181]	0.224 [0.220,0.228]	0.064 [0.058,0.070]	0.092 [0.079,0.105]
SNF days	0.198 [0.197,0.200]	0.217 [0.215,0.219]	0.184 ^A [0.182,0.186]	0.152 [0.147,0.158]	0.176 ^A [0.166,0.187]
Outpatient days	0.568 [0.566,0.570]	0.624 [0.620,0.627]	0.494 [0.491,0.497]	0.702 ^A [0.691,0.712]	0.695 ^A [0.673,0.717]
Primary care days	0.347 [0.347,0.348]	0.348 [0.347,0.349]	0.356 [0.355,0.357]	0.258 [0.256,0.261]	0.327 [0.321,0.332]
Specialist days	0.361 [0.360,0.361]	0.334 [0.333,0.335]	0.369 [0.368,0.370]	0.527 [0.523,0.531]	0.464 [0.457,0.471]
Carrier days ²	1.435 [1.433,1.437]	1.410 ^A [1.407,1.413]	1.410 ^A [1.407,1.413]	1.860 [1.849,1.871]	1.739 [1.719,1.760]
ED visits	0.055 [0.055,0.055]	0.054 [0.054,0.054]	0.055 [0.055,0.055]	0.064 [0.063,0.065]	0.071 [0.069,0.072]
PCT ED visits	0.008 [0.008,0.008]	0.008 [0.008,0.008]	0.007 [0.007,0.007]	0.010 [0.010,0.010]	0.011 [0.010,0.011]

Observations: n=1,483,567 unique beneficiaries, N=8,892,882 bene/season/locations

¹ Confidence intervals for ratio estimates using linearized standard errors

² Days of physician care in carrier file; includes physician visits/services in any setting

^A Estimates not significantly different from each other at p=0.01 (within row)

Estimated incidence rate ratios from the population-averaged negative binomial

regression models are presented in Table 4.3 for the key variables. Full results are presented in Appendix F. Models were adjusted for location, age, season, race, sex, Medicaid status, and comorbidities using quintiles of the Hierarchical Condition Category community score.

The main effect estimates for migration categories suggest that after adjusting for covariates, Snowbirds have higher incidence rates of primary care treatable ED visits, total ED visits, and specialist visits, but fewer SNF, inpatient, and hospice days per month of healthcare received than people who received care in the North only.

Table 4.3: Use of Healthcare Services by Migration Status

	(1)	(2)	(3)	(4)	(5)	(6)
	PCT ED visits	ED visits	Inpt days	Specialist visits	SNF days	Hospice days
Migration categories						
N only	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]
S only	0.896*** [0.852,0.943]	0.980 [0.957,1.003]	1.090*** [1.049,1.131]	1.138*** [1.124,1.152]	0.896* [0.819,0.980]	1.885*** [1.699,2.091]
Snowbird	1.669*** [1.589,1.753]	1.278*** [1.245,1.312]	0.721*** [0.682,0.762]	1.331*** [1.313,1.350]	0.338*** [0.283,0.404]	0.130*** [0.110,0.155]
Other	2.059*** [1.913,2.216]	1.813*** [1.746,1.883]	1.024 [0.944,1.112]	1.124*** [1.099,1.150]	0.387*** [0.296,0.505]	0.261*** [0.206,0.331]
PC at loc	0.711*** [0.693,0.729]	0.844*** [0.834,0.855]	0.864*** [0.847,0.881]	1.180*** [1.171,1.188]	2.010*** [1.918,2.106]	0.346*** [0.330,0.363]
Migration category X PC at location						
S only X PC at loc	0.960* [0.926,0.996]	0.912*** [0.897,0.928]	0.914*** [0.889,0.939]	0.999 [0.989,1.009]	1.038 [0.964,1.116]	0.730*** [0.684,0.780]
Snowbird X PC at loc	0.590*** [0.559,0.623]	0.702*** [0.683,0.721]	1.185*** [1.122,1.251]	0.938*** [0.924,0.951]	2.047*** [1.721,2.436]	3.105*** [2.673,3.607]
Other X PC at loc	0.495*** [0.454,0.540]	0.545*** [0.523,0.569]	0.939 [0.862,1.024]	1.096*** [1.070,1.123]	2.231*** [1.712,2.907]	2.471*** [2.040,2.992]

Exponentiated coefficients; 95% confidence intervals in brackets

X denotes interaction terms

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

However, the main effect for migration category doesn't tell the whole story. Having local primary care strongly reduced rates of both PCT ED visits and total ED visits for all migration categories, as shown in the main and interaction effects of the local primary care variable. To ease interpretation of these interactions, we provide predicted rates of ED visits and PCT ED visits per month of healthcare received in Table 4.4. This table also includes a within-migration category percent reduction for the effect of having local primary care. The predicted rate of primary care treatable ED visits for people in the North only migration category who have local primary care is 0.012 PCT ED visits per month of healthcare received, while the predicted rate for the same people with no local primary care is 0.017 PCT ED visits per month of healthcare received—a 29% reduction associated with having received local primary care. For Snowbirds, this effect is much larger: a 58% reduction in primary care treatable ED visits associated with having received local primary care. Note, too, that the percentage reduction of local primary care on primary care treatable ED visits associated with having local primary care is larger than the percent reduction for total ED visits for all migration categories; this may suggest that primary care is best at reducing ED visits for minor problems, but doesn't necessarily improve health enough to reduce more serious ED visits.

Table 4.4: Predictive margins: incidence rates of ED visits per month of healthcare received

Migration Cat.	Local primary care	PCT ED visits		All ED visits	
		Rate	95% CI	Rate	95% CI
North only	No	0.017	[0.017,0.018]	0.101	[0.099,0.102]
	Yes	0.012	[0.012,0.012]	0.085	[0.085,0.085]
	Difference	-29%		-16%	
South only	No	0.017	[0.017,0.017]	0.112	[0.111,0.114]
	Yes	0.012	[0.012,0.012]	0.087	[0.086,0.087]
	Difference	-32%		-23%	
Snowbird	No	0.028	[0.026,0.029]	0.133	[0.130,0.136]
	Yes	0.012	[0.011,0.012]	0.079	[0.078,0.080]
	Difference	-58%		-41%	
Other	No	0.037	[0.034,0.039]	0.198	[0.191,0.205]
	Yes	0.013	[0.012,0.014]	0.091	[0.089,0.093]
	Difference	-65%		-54%	

Marginal effect calculation uses migration categories as observed, varying only the local primary care variable. PCT=Primary Care Treatable

Discussion

Our analysis identified about 5.5% of the Medicare aged FFS population as seasonal migrators. Most of those migrators followed the traditional Snowbird pattern of migration, spending warm months in the north and cold months in the south. There are no good national level estimates of prevalence of seasonal migration among elderly people to compare (Siegel 2011; Smith and House 2006). Smith and House (2006) estimate 4.7% of the Florida population aged 55 and over are temporary residents, staying more than a month, but reporting that Florida is not their usual place of residence. However, this definition is sufficiently different from ours (by including those aged 55 to 64 and requiring that people say they don't "usually" live in Florida) that it is difficult to compare. Estimates in Hogan (1987) based on a special 1980 Census question suggest that about 1.1% of people aged 65 and older were non-permanent residents of either Arizona or Florida as of April first. However, this estimate is noted to be low, as many Snowbirds would have returned to the north by

April. Furthermore, the current group of people aged 65 and older is of a different generation than those aged 65 and older in 1980 and may have very different patterns of seasonal migration.

Given these other estimates, our estimate of 4.11% may be a bit high, likely including people who received healthcare while on vacation, rather than true temporary residents. However, it is interesting to note that the proportion of Snowbirds with primary care in the north is only about five percentage points higher than the proportion of Snowbirds with primary care in the south (74.4% vs. 69.6%). We might consider that those Snowbirds with primary care in both the north and south are almost certainly seasonal migrators (48.7% of Snowbirds), giving a lower bound estimate of about 2% of the US Medicare aged FFS population being true Snowbird seasonal migrators.

The estimated rates of healthcare service use by Snowbirds are a new contribution to the literature on seasonal migration, as is the estimate of the association between having local primary care and rates of primary care treatable ED visits. Our data suggest that Snowbirds rates of primary care treatable ED visits are reduced more than 50 percent when they have local primary care.

Limitations

Because it relies on claims data to determine where people live, this analysis has several limitations. Although we have relatively complete coverage by season, with at least one observation for 81% of all beneficiary–seasons eligible for the analysis, we do not know where beneficiaries are in between observations. Some beneficiaries will have very limited location information if they rarely seek medical care or if they seek medical care mostly outside the Medicare payment system. We also must exclude people on Medicare Advantage plans, which is a significant proportion of Medicare

beneficiaries (25% in 2011: Jacobson et al. 2015). As a result, our analysis doubtless misses many seasonal migrators who didn't need medical care while away and may mislabel as migrators others who received medical care while on vacation.

Still, in the absence of any other information about the healthcare service use of this population, this analysis should yield insight about the population we are able to track relatively well using claims. Our findings may encourage further research to identify and understand elderly seasonal migrators.

Conclusions

As ACOs and other new delivery models face increased provider responsibility for paying for all healthcare their members use, regardless of their location, it is more relevant than ever to understand the healthcare service use of seasonal migrators. Our results suggest that having a local primary care physician is associated with significantly fewer primary care treatable ED visits, as well as a smaller reduction in total ED visits. ACOs may want to consider programs that encourage their members to seek primary care in their temporary residence locations, as well as programs to coordinate care outside their geographic area.

Chapter 5

Conclusions

The purpose of this research was to create an alternative to the Billings/NYU algorithm that would solve some of the implementation problems users have faced with that algorithm and to test another approach to identifying less urgent ED visits in a Medicare population. We have succeeded in these aims. The new Minnesota algorithm is as good a predictor of death and hospitalizations as the validated Ballard version of the Billings algorithm. And unlike the Billings algorithm, it will not be affected by the change to ICD10; furthermore, because it doesn't rely on diagnoses, it can be used to distinguish severity among visits by people who have the same chronic condition. Although we have tested it on a Medicare sample, it was also designed to be applicable to any US population. By contrast, the Billings algorithm, which was developed in an urban, low-income, primary-care deprived population, may have heavily influenced the severity assigned to the diagnoses used in the algorithm. The Medicare population is probably among the worst served by the Billings algorithm, since the clinical severity of a diagnosis can be very different for an elderly population compared to an all-ages population like that used in creating the Billings algorithm. Future work can validate the new algorithm in broader populations.

Chapter two of this work describes the design of the Minnesota algorithm and validates it in a Medicare population using death and hospitalization as outcomes.

The results are compared to the Ballard version of the Billings algorithm and the performance of the new algorithm found to be as good or better than the older one.

Chapters three and four of this work are applications of the new algorithm. Chapter three asks whether increased use of primary care reduces primary care treatable ED visits, while Chapter four explores the primary care seeking behavior of elderly seasonal migrators and tests whether having local primary care reduces primary care treatable ED visits. Both analyses provide evidence of an association between having a primary care relationship and a reduction in primary care treatable ED visits, though causality is difficult to infer because we have only observational data. Taken together, the two chapters form a compelling argument for some degree of causality, nonetheless. Part one of Chapter three uses geographic variation in how often people visit a primary care physician while controlling for demographic and health differences across those geographies; part two of Chapter three uses individual variability in preference for having a primary care relationship to identify an effect while again controlling for demographic and health differences. A bounded treatment effect analysis suggests that a positive causal effect for primary care is possible under a range of assumptions about the treatment selection process and error in determining treatment status (i.e., whether an individual has a primary care relationship). Two scenarios identified a positive treatment effect for certain; all other realistic scenarios suggested at worst a small negative effect and at best a very strong positive effect. Chapter four uses seasonal migrators as their own control group to determine a treatment effect of having local primary care, looking at the rate of primary care treatable ED visits when migrators are near their primary care physician and when they're away. Again, a positive treatment effect is found.

This work also contributes to analysis methods, by developing Hospital Referral Region (HRR) groups that can be used with county-level residence or covariate data. Previously, to use HRRs in analysis, the researcher needed zip code level residence

and covariate data, which can be difficult to find. For example, the rich local data available in the Area Health Resource File (AHRF) is available at the county, but not the zip code level.

Another analytic method developed in this work is the use of healthcare claims data to identify seasonal migrators, a group about whom relatively little is known. The trend of payment programs that require providers to manage the healthcare services used by their patients, regardless of location, makes this a timely contribution.

As discussed in the body of this work, there are some key limitations. The validation exercise in Chapter two suffers from a lack of a gold standard. We were only able to validate the measure using high severity outcomes of death and hospitalization. The ideal validation would be to compare algorithm categorization of visits to some objective analysis of the primary care treatability of the visits.

Another limitation of this work, as noted above, is that analyses in Chapters three and four are unable to definitively identify a positive, causal treatment effect of primary on reducing primary care treatable ED visits. Identifying a robust association is an incremental addition to the existing literature on this relationship.

Future work

The next steps for this research will likely include further validation of the measure, both in broader populations and against better standards. One possible procedure to get closer to a gold standard comparison might involve asking both primary care and ED physicians to review medical records from ED visits and assess their primary care treatability, as Billings et al. did in their original analysis. This would likely be expensive and time-consuming, but would probably be the preferred approach. A less resource-intensive process would be to ask ED physicians to rate the primary care treatability of all patients they see for some period, then compare physician ratings

to algorithm ratings.

Another line of inquiry might include implementing the new measure in existing data from experimental and quasi-experimental interventions where the Billings algorithm has been used. The Oregon Health Insurance experiment group did not find a change in the severity of ED visits for people who were previously uninsured, but it may be that the Billings algorithm is not sufficiently sensitive to detect a change. Changing the severity of ED visits as measured with the Billings algorithm requires changing the diagnoses in the population using the ED. Since many ED visits are caused by chronic conditions, and even more visits are by people who have chronic conditions (whether or not that condition is what brings them to the ED), the Billings algorithm approach of using diagnoses to rate visit severity may be unable to detect the kind of smaller, short-term changes we would expect to see from an experiment or quasi-experiment.

Policy implications

The new measure will be most useful in evaluating the effect of interventions to improve access to and use of non-emergency care. It is not suitable for determining whether to pay for an individual ED visit, for example, as in a plan passed by the Washington legislature but never enacted, that would have refused payment for primary care treatable ED visits by Medicaid beneficiaries (WSHA 2012). Individuals deciding where to present for treatment are not qualified to determine the severity of their signs and symptoms. Programs to reduce unnecessary ED visits should focus on access to decision-making tools like nurse lines to help people decide whether the ED is the best venue for care, same-day and next-day primary care appointments, and extended office hours: in short, making it easier for people to access the appropriate care venue in a timely way.

We hope that other researchers will find the new Minnesota algorithm to be a useful addition to scholarship on the efficient use of healthcare resources.

Bibliography

- AHRQ Quality Indicators—Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions*. 2001. Rockville, MD: Agency for Healthcare Research and Quality. <http://www.ahrq.gov/downloads/pub/ahrqqi/pqiguide.pdf>.
- American College of Emergency Physicians [ACEP]. 2013. “Ten Things Physicians and Patients Should Question.” October 14. Accessed April 2, 2015. <http://www.choosingwisely.org/doctor-patient-lists/american-college-of-emergency-physicians/>.
- Bacon-Shone, John. 2011. “A Short History of Compositional Data Analysis.” In *Compositional Data Analysis: Theory and Applications*, edited by Vera Pawlowsky-Glahn and Antonella Buccianti. Chichester, UK: John Wiley & Sons, Ltd.
- Ballard, Dustin W, Mary Price, Vicki Fung, Richard Brand, Mary E Reed, Bruce Fireman, Joseph P Newhouse, Joseph V Selby, and John Hsu. 2010. “Validation of an algorithm for categorizing the severity of hospital emergency department visits.” *Medical care* 48, no. 1 (January): 58–63.
- Barnett, Karen, Stewart W Mercer, Michael Norbury, Graham Watt, Sally Wyke, and Bruce Guthrie. 2012. “Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study.” *Lancet* 380, no. 9836 (July): 37–43.
- Begley, Charles E, Rachel Westheimer Vojvodic, Munseok Seo, and Keith Burau. 2006. “Emergency Room Use and Access to Primary Care: Evidence from Houston, Texas.” *Journal of health care for the poor and underserved* 17 (3): 610–624.
- Belotti, Federico, Gordon Hughes, and Andrea Piano Mortari. 2013. “xsmle: A Stata Command for spatial panel-data models estimation.” <http://econpapers.repec.org/paper/bocdsug13/09.htm>.

- Bentley, Patricia N, Andrew G Wilson, Mary Elizabeth Derwin, Robbie Scodellaro, and Raymond E Jackson. 2002. "Reliability of assigning correct current procedural terminology—4 E/M codes." *Annals of Emergency Medicine* 40, no. 3 (September): 269–274.
- Bezzina, Andrew J, Peter B Smith, David Cromwell, and Kathy Eagar. 2005. "Primary care patients in the emergency department: who are they? A review of the definition of the 'primary care patient' in the emergency department." *Emergency medicine Australasia : EMA* 17, nos. 5-6 (October): 472–479.
- Billings, John. 2012. "NYU ED algorithm background." Accessed December 28. <http://wagner.nyu.edu/faculty/billings/nyued-background.php>.
- Billings, John, Nina Parikh, and Tod Mijanovich. 2000a. "Emergency department use in New York City: a substitute for primary care?" *Issue brief (Commonwealth Fund)*, no. 433:1–5.
- . 2000b. "Emergency department use in New York City: a survey of Bronx patients." *Issue brief (Commonwealth Fund)*, no. 435 (November): 1–5.
- . 2000c. "Emergency Room Use: The New York Story." *Issue brief (Commonwealth Fund)*, no. 434 (November): 1–11.
- Billings, John, Lisa Zeitel, Joanne Lukomnik, Timothy S Carey, Arthur E Blank, and Laurie Newman. 1993. "Impact of socioeconomic status on hospital use in New York City." *Health affairs (Project Hope)* 12 (1): 162–173.
- Birnbaumer, Diane M. 2014. "The Elder Patient." In *Rosen's Emergency Medicine*, 8th ed., edited by John A Marx, Robert S Hockberger, and Ron M Walls, 2351–2355. Philadelphia: Saunders.
- Brunt, Christopher S. 2011. "CPT fee differentials and visit upcoding under Medicare Part B." *Health Economics* 20, no. 7 (April): 831–841.
- Buczko, William. 1992. "Factors affecting interstate use of inpatient care by Medicare beneficiaries." *Health Services Research* 27, no. 3 (August): 295–315.
- . 1994. "Differences in hospitalizations among seasonal migrants, adjacent-state and in-state aged Medicare beneficiaries." *Medical care* 32, no. 3 (March): 308–314.
- Bureau of Economic Analysis [BEA]. 2015. "BEA Regions." Accessed June 17. <https://www.bea.gov/regional/docs/regions.cfm>.

- Centers for Medicare & Medicaid Services [CMS]. 2012a. *Medicare Claims Processing Manual [Pub 100-04]*. Accessed May 5. <http://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/clm104c12.pdf>.
- . 2012b. *Medicare Claims Processing Manual [Pub 100-04]*. Accessed May 5. <http://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/clm104c4.pdf>.
- . 2014. “Standard Analytic Files — LDS.” Accessed December 16. <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/StandardAnalyticalFiles.html>.
- . 2011. *Medicare National Coverage Determinations Manual*. February. <http://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Internet-Only-Manuals-IOMs-Items/CMS014961.html>.
- . 2012c. “Appendices Medicare Fee-for-Service 2011 Improper Payments Report.” October 2. Accessed February 1, 2015. <http://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/Medicare-FFS-Compliance-Programs/CERT/CERT-Reports-Items/Nov2011Appendix.html>.
- . 2012d. “Comprehensive Primary Care (CPC) Initiative: Primary Care Practice Solicitation.” June 13. Accessed April 8, 2015. http://innovation.cms.gov/Files/x/CPC_PracticeSolicitation.pdf.
- . 2015. *Medicare Geographic Variation Public Use File*. http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Geographic-Variation/GV_PUF.html.
- Chang, Chiang-Hua, Therese A Stukel, Ann Barry Flood, and David C Goodman. 2011. “Primary care physician workforce and Medicare beneficiaries’ health outcomes.” *JAMA: the journal of the American Medical Association* 305 (20): 2096.
- Chen, Ronald C, Trevor J Royce, Martine Extermann, and Bryce B Reeve. 2012. “Impact of Age and Comorbidity on Treatment and Outcomes in Elderly Cancer Patients.” *Seminars in Radiation Oncology* 22, no. 4 (October): 265–271.
- Coleman, Eric A, Theresa B Eilertsen, Andrew M Kramer, David J Magid, Arne Beck, and Doug Conner. 2001. “Reducing emergency visits in older adults with chronic illness. A randomized, controlled trial of group visits.” *Effective clinical practice* 4, no. 2 (March): 49–57.

- Collard, Rose M, Han Boter, Robert A Schoevers, and Richard C Oude Voshaar. 2012. "Prevalence of Frailty in Community-Dwelling Older Persons: A Systematic Review." *Journal of the American Geriatrics Society* 60, no. 8 (August): 1487–1492.
- Counsell, Steven R, Christopher M Callahan, Daniel O Clark, Wanzhu Tu, Anna B Buttar, Timothy E Stump, and Gretchen D Ricketts. 2007. "Geriatric care management for low-income seniors: a randomized controlled trial." *JAMA: the journal of the American Medical Association* 298, no. 22 (December): 2623–2633.
- D'Arcy, Laura P, Sally C Stearns, Marisa E Domino, Laura C Hanson, and Morris Weinberger. 2012. "Is Geriatric Care Associated with Less Emergency Department Use?" *Journal of the American Geriatrics Society* 61, no. 1 (December): 4–11.
- Dartmouth Institute for Health Policy & Clinical Practice. 2015. "The Dartmouth Atlas of Health Care." <http://www.dartmouthatlas.org>.
- Davis, James W, Ronald Y Fujimoto, Henry Chan, and Deborah T Juarez. 2010. "Identifying characteristics of patients with low urgency emergency department visits in a managed care setting." *Managed care (Langhorne, Pa.)* 19, no. 10 (October): 38–44.
- Delmarva Foundation. 2008. *Reducing Avoidable Emergency Room Visits Baseline Report: Report to California Department of Health Care Services Medi-Cal Managed Care Division*. August. http://www.dhcs.ca.gov/dataandstats/reports/Documents/MMCD_Qual_Rpts/EQRO_QIPs/ER%20Collaborative%202008%20FINAL.pdf.
- Esri. 2012. *HRR boundaries for AGO*. Modified version of shapefile created by Dartmouth Atlas. Link to original data and description of modifications available at ArcGIS URL provided here. <http://www.arcgis.com/home/item.html?id=00bc657ae97a424b92f7f9c38a178503>.
- Felton, Brent M, Earl J Reisdorff, Christopher N Krone, and Gus A Laskaris. 2011. "Emergency department overcrowding and inpatient boarding: a statewide glimpse in time." *Academic emergency medicine* 18, no. 12 (December): 1386–1391.
- Geruso, Michael, and Timothy Layton. 2014. "Risk Selection, Risk Adjustment, and Manipulable Medical Coding: Evidence from Medicare." Working paper, January 20. <https://sites.google.com/site/timothyjlayton/UpcodingMA.pdf>.

- Gilboy, Nicki, Paula Tanabe, Debbie A Travers, and Alexander Rosenau. 2011. *Emergency Severity Index (ESI), A Triage Tool for Emergency Department Care, Version 4, Implementation Handbook, 2012 Edition*. Rockville, MD: Agency for Healthcare Research / Quality (AHRQ). <http://www.ahrq.gov/research/esi/esihandbk.pdf>.
- Gordon, James A, John Billings, Brent R Asplin, and Karin V Rhodes. 2001. "Safety net research in emergency medicine proceedings of the Academic Emergency Medicine Consensus Conference on 'The Unraveling Safety Net'." *Academic emergency medicine* 8, no. 11 (November): 1024–1029.
- Greene, Jan. 2014. "Price Transparency in the Emergency Department." *Annals of Emergency Medicine* 64, no. 1 (July): A13–A15.
- Hampton, Tracy. 2008. "Experts predict visits by baby boomers will soon strain emergency departments." *JAMA: the journal of the American Medical Association* 299, no. 22 (June): 2613–2614.
- Al-Haque, Shahed, Mehmet Erkan Ceyhan, Stephanie H Chan, and Deborah J Nightingale. 2015. "Responding to Traveling Patients' Seasonal Demand for Health Care Services." *Military Medicine* 180, no. 1 (January): 111–117.
- Hastings, Susan, Karen Stechuchak, Eugene Oddone, Morris Weinberger, Dana Tucker, William Knaack, and Kenneth Schmader. 2012. "Older veterans and emergency department discharge information." *BMJ Quality & Safety* 21, no. 10 (September): 835–842.
- Health Resources and Services Administration [HRSA]. 2014. *Area Health Resource File*. <http://ahrf.hrsa.gov/download.htm>.
- Healthcare Bluebook. 2014. *Office Visit, New Patient, Level 1*. https://healthcarebluebook.com/page_ProcedureDetails.aspx?id=218&dataset=MD&g=Office+Visit,+New+Patient,+Level+1.
- Hogan, Timothy D. 1987. "Determinants of the seasonal migration of the elderly to Sunbelt states." *Research on Aging* 9, no. 1 (March): 115–133.
- Hunold, Katherine M, Natalie L Richmond, Anna E Waller, Malcolm P Cutchin, Paul R Voss, and Timothy F Platts-Mills. 2014. "Primary Care Availability and Emergency Department Use by Older Adults: A Population-Based Analysis." *Journal of the American Geriatrics Society* 62, no. 9 (August): 1699–1706.

- Hvidt, Lisa, Kristian Nebelin Hvidt, Kim Madsen, and Thomas Andersen Schmidt. 2014. "Comprehension deficits among older patients in a quick diagnostic unit." *Clinical Interventions in Aging* (April): 705–6.
- Institute of Medicine [IOM] Committee on the Future of Emergency Care in the United States Health System. 2007. *Hospital-Based Emergency Care: At the Breaking Point*. National Academy Press.
- Jacobson, Gretchen, Anthony Damico, Tricia Neuman, and Marcia Gold. 2015. "Medicare Advantage 2015 Spotlight: Enrollment Market Update." June 30. Accessed August 21, 2015. <http://kff.org/medicare/issue-brief/medicare-advantage-2015-spotlight-enrollment-market-update/>.
- Jones, Kari, Hannah Paxton, Reidar Hagtvedt, and Jeff Etchason. 2011. "An Analysis of the New York University Emergency Department Algorithm's Suitability for Use in Gauging Changes in ED Usage Patterns." *Medical care* (December): 1–10.
- Kane, Robert L, Joseph G Ouslander, Itamar B Abrass, and Barbara Resnick. 2013. *Essentials of Clinical Geriatrics*. 7th ed. New York: McGraw–Hill Education.
- Kaskie, Brian, Maksym Obrizan, Elizabeth A Cook, Michael P Jones, Li Liu, Suzanne Bentler, Robert B Wallace, John F Geweke, Kara B Wright, Elizabeth A Chrischilles, Claire E Pavlik, Robert L Ohsfeldt, Gary E Rosenthal, and Fredric D Wolinsky. 2010. "Defining emergency department episodes by severity and intensity: A 15-year study of Medicare beneficiaries." *BMC Health Services Research* 10 (1): 173.
- Kaskie, Brian, Maksym Obrizan, Michael P Jones, Suzanne Bentler, Paula Weigel, Jason Hockenberry, Robert B Wallace, Robert L Ohsfeldt, Gary E Rosenthal, and Fredric D Wolinsky. 2011. "Older adults who persistently present to the emergency department with severe, non-severe, and indeterminate episode patterns." *BMC Geriatrics* 11, no. 1 (October): 65.
- Kellermann, Arthur L, and Robin M Weinick. 2012. "Emergency departments, Medicaid costs, and access to primary care—understanding the link." *The New England journal of medicine* 366, no. 23 (June): 2141–2143.
- LaMantia, Michael A, Timothy F Platts-Mills, Kevin Biese, Christine Khandelwal, Cory Forbach, Charles B Cairns, Jan Busby-Whitehead, and John S Kizer. 2010. "Predicting hospital admission and returns to the emergency department for elderly patients." *Academic emergency medicine* 17, no. 3 (March): 252–259.

- Landis, J Richard, and Gary G Koch. 1977. "The measurement of observer agreement for categorical data." *Biometrics* 33, no. 1 (March): 159–174.
- Lowe, Robert A, and Rongwei Fu. 2008. "Can the Emergency Department Algorithm Detect Changes in Access to Care?" *Academic emergency medicine* 15, no. 6 (June): 506–516.
- Lowe, Robert A, A Russell Localio, Donald F Schwarz, Sankey Williams, Lucy Wolf Tuton, Staci Maroney, David Nicklin, Neil Goldfarb, Deneen D Vojta, and Harold I Feldman. 2005. "Association between primary care practice characteristics and emergency department use in a Medicaid managed care organization." *Medical care* 43, no. 8 (August): 792–800.
- Lundgren, Rachel S, C Bradley Kramer, Frederick P Rivara, Jin Wang, David M Heimbach, Nicole S Gibran, and Matthew B Klein. 2009. "Influence of Comorbidities and Age on Outcome Following Burn Injury in Older Adults." *Journal of Burn Care & Research* 30, no. 2 (March): 307–314.
- Martín-Fernández, Josep Antoni, Javier Palarea-Albaladejo, and Ricardo Antonio Olea. 2011. "Dealing with zeros." In *Compositional Data Analysis: Theory and Applications*, edited by Vera Pawlowsky-Glahn and Antonella Buccianti. Chichester, UK: John Wiley & Sons, Ltd.
- McCarthy, Ian, Daniel L. Millimet, and Manan Roy. 2015. "Bounding treatment effects: A command for the partial identification of the average treatment effect with endogenous and misreported treatment assignment." *The Stata Journal* 15 (2): 411–436.
- McClelland, Mark, Brent Asplin, Stephen K Epstein, Keith Eric Kocher, Randy Pilgrim, Jesse Pines, Elaine Judith Rabin, and Niels Kumar Rathlev. 2014. "The Affordable Care Act and emergency care." *American Journal of Public Health* 104, no. 10 (October): e8–10.
- MedPAC. 2012. *Health Care Spending and the Medicare Program*. Washington: MedPAC, June. <http://www.medpac.gov/documents/Jun12DataBookEntireReport.pdf>.
- Mello, Michelle M, Sally C Stearns, Edward C Norton, and Thomas C Ricketts. 2003. "Understanding biased selection in Medicare HMOs." *Health Services Research* 38, no. 3 (June): 961–992.

- Musso, Mandi W, J Nelson Perret, Taylor Sanders, Ross Daray, Kyle Anderson, Melissa Lancaster, David Lim, and Glenn N Jones. 2015. "Patients' Comprehension of Their Emergency Department Encounter: A Pilot Study Using Physician Observers." *Annals of Emergency Medicine* 65, no. 2 (February): 151–155.e4.
- National Center for Health Statistics [NCHS]. 2008. *2008 National Hospital Ambulatory Medical Care Survey Public Use Data File Documentation*. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/.
- . 2009. *2009 National Hospital Ambulatory Medical Care Survey Public Use Data File Documentation*. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/.
- . 2014. "Health, United States, 2013." Accessed March 12, 2015. <http://www.cdc.gov/nchs/data/hus/hus13.pdf>.
- Office of Evaluation and Inspections [OIG OEI]. 2014. *Improper Payments for Evaluation and Management Services Cost Medicare Billions in 2010*. Government report OEI-04-10-00181. Office of the Inspector General, Department of Health and Human Services, May. <https://oig.hhs.gov/oei/reports/oei-04-10-00181.pdf>.
- Olshaker, Jonathan S. 2009. "Managing emergency department overcrowding." *Emergency Medicine Clinics of NA* 27, no. 4 (November): 593–603.
- Piccirillo, Jay F, Anna Vlahiotis, Laurel B Barrett, Kellie L Flood, Edward L Spitznagel, and Ewout W Steyerberg. 2008. "The changing prevalence of comorbidity across the age spectrum." *Critical Reviews in Oncology/Hematology* 67, no. 2 (August): 124–132.
- Piehl, Mark D, Conrad J Clemens, and Jerry D Joines. 2000. "'Narrowing the gap': decreasing emergency department use by children enrolled in the Medicaid program by improving access to primary care." *Arch Pediatr Adolesc Med*, no. 8:791–795.
- Pitts, Stephen R, Emily R Carrier, Eugene C Rich, and Arthur L Kellermann. 2010. "Where Americans Get Acute Care: Increasingly, It's Not At Their Doctor's Office." *Health Affairs* 29, no. 9 (September): 1620–1629.

- Pukurdpol, Paul, Jennifer L Wiler, Renee Y Hsia, and Adit A Ginde. 2014. "Association of Medicare and Medicaid Insurance With Increasing Primary Care-treatable Emergency Department Visits in the United States." *Academic emergency medicine* 21, no. 10 (October): 1135–1142.
- Rich, Michael W. 2005. "Heart Failure in the Oldest Patients: The Impact of Comorbid Conditions." *The American Journal of Geriatric Cardiology* 14, no. 3 (May): 134–141.
- Roberts, David C, Mary Pat McKay, and Alison Shaffer. 2008. "Increasing Rates of Emergency Department Visits for Elderly Patients in the United States, 1993 to 2003." *Annals of Emergency Medicine* 51, no. 6 (June): 769–774.
- Roth, Jean. 2014. *CMS' National Plan and Provider Enumeration System (NPPES) Files in SAS, Stata, and CSV format*. <http://www.nber.org/data/npi.html>.
- . 2015. *County Distance Database*. <http://www.nber.org/data/county-distance-database.html>.
- Ruger, Jennifer Prah, Lawrence M Lewis, and Christopher J Richter. 2012. "Patterns and factors associated with intensive use of ED services: implications for allocating resources." *American Journal of Emergency Medicine* (July): 1–11.
- Schiff, Gordon D. 2011. "System dynamics and dysfunctionalities: levers for overcoming emergency department overcrowding." *Academic emergency medicine* 18, no. 12 (December): 1255–1261.
- Schulte, Fred, and David Donald. 2012. "How doctors and hospitals have collected billions in questionable Medicare fees: Center investigation suggests costs from upcoding and other abuses likely top \$11 billion." In "Medicare: Cracking the Codes," *The Center for Public Integrity* (September 15). <http://www.publicintegrity.org/2012/09/15/10810/how-doctors-and-hospitals-have-collected-billions-questionable-medicare-fees>.
- Scott, John C, Douglas A Conner, Ingrid Venohr, Glenn Gade, Marlene McKenzie, Andrew M Kramer, Lucinda Bryant, and Arne Beck. 2004. "Effectiveness of a group outpatient visit model for chronically ill older health maintenance organization members: a 2-year randomized trial of the cooperative health care clinic." *Journal of the American Geriatrics Society* 52, no. 9 (September): 1463–1470.

- Shankar, Kalpana N, Bhavnit K Bhatia, and Jeremiah D Schuur. 2014. "Toward Patient-Centered Care: A Systematic Review of Older Adults' Views of Quality Emergency Care." *Annals of Emergency Medicine* 63, no. 5 (May): 529–550.e1.
- Siegel, Jacob S. 2011. *The Demography and Epidemiology of Human Health and Aging*. Dordrecht, The Netherlands: Springer Science & Business Media, September.
- Smith, Stanley K, and Mark House. 2006. "Snowbirds, sunbirds, and stayers: seasonal migration of elderly adults in Florida." *The journals of gerontology. Series B, Psychological sciences and social sciences* 61, no. 5 (September): S232–9.
- Song, Yunjie, Jonathan Skinner, Julie Bynum, Jason Sutherland, John E Wennberg, and Elliott S Fisher. 2010. "Regional Variations in Diagnostic Practices." *The New England journal of medicine* 363, no. 1 (July): 45–53.
- StataCorp. 2015a. *Stata Statistical Software*. Release 14. College Station, TX: StataCorp LP.
- . 2015b. *Stata Structural Equation Modeling Reference Manual*. Release 14. College Station, TX: StataCorp LP.
- . 2015c. *Stata Treatment-Effects Reference Manual: potential outcomes/counterfactual outcomes*. Release 14. College Station, TX: StataCorp LP.
- Stedman's Medical Dictionary—27th edition*. 2000. Philadelphia: Lippincott Williams & Wilkins.
- Terza, Joseph V. 1998. "Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects." *Journal of Econometrics* 84:129–154.
- Washington State Hospital Association [WSHA]. 2011. *Potentially Avoidable Emergency Room Use*. Report to Washington State Hospital Association. February. <http://www.wsha.org/files/127/ERReport2.pdf>.
- . 2012. *Background Information: Medicaid Emergency Department Use*, April 16. <http://www.wsha.org/files/ER%20Background%20Information%20April%202016.pdf>.
- Weinick, Robin M, Rachel M Burns, and Ateev Mehrotra. 2010. "Many Emergency Department Visits Could Be Managed At Urgent Care Centers And Retail Clinics." *Health Affairs* 29, no. 9 (September): 1630–1636.

- Werner, Rachel M, Anne Canamucio, Steven C Marcus, and Christian Terwiesch. 2014. "Primary Care Access and Emergency Room Use Among Older Veterans." *Journal of General Internal Medicine* 29, no. S2 (April): 689–694.
- Wharam, J Frank, Bruce E Landon, Alison A Galbraith, Ken P Kleinman, Stephen B Soumerai, and Dennis Ross-Degnan. 2007. "Emergency department use and subsequent hospitalizations among members of a high-deductible health plan." *JAMA: the journal of the American Medical Association* 297, no. 10 (March): 1093–1102.
- Wolinsky, Fredric D, Li Liu, Thomas R Miller, Hyonggin An, John F Geweke, Brian Kaskie, Kara B Wright, Elizabeth A Chrischilles, Claire E Pavlik, Elizabeth A Cook, Robert L Ohsfeldt, Kelly K Richardson, Gary E Rosenthal, and Robert B Wallace. 2008. "Emergency department utilization patterns among older adults." *The journals of gerontology. Series A, Biological sciences and medical sciences* 63, no. 2 (February): 204–209.
- Yancik, R, M N Wesley, L A Ries, R J Havlik, S Long, B K Edwards, and J W Yates. 1998. "Comorbidity and age as predictors of risk for early mortality of male and female colon carcinoma patients: a population-based study." *Cancer* 82, no. 11 (June): 2123–2134.

Appendix A

ED indicator procedures

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
Drugs administered: antianxiety	
J2060	Injection, lorazepam, 2 mg
J3360	Injection, diazepam, up to 5 mg
Drugs administered: antiarrhythmic	
J0150	Injection, adenosine for therapeutic use, 6 mg
J0282	Injection, amiodarone hydrochloride, 30 mg
Drugs administered: antibiotic and antifungal	
J3370	Injection, vancomycin hcl, 500 mg
J0295	Injection, ampicillin sodium/sulbactam sodium, per 1.5 gm
J0456	Injection, azithromycin, 500 mg
J0690	Injection, cefazolin sodium, 500 mg
J0713	Injection, ceftazidime, per 500 mg
J0743	Injection, cilastatin sodium; imipenem, per 250 mg
J0744	Injection, ciprofloxacin for intravenous infusion, 200 mg
J1956	Injection, levofloxacin, 250 mg
J2020	Injection, linezolid, 200mg
J2185	Injection, meropenem, 100 mg
J2280	Injection, moxifloxacin, 100 mg
J2543	Injection, piperacillin sodium/tazobactam sodium, 1 gram/0.125 grams
J1450	Injection fluconazole, 200 mg
Drugs administered: anticoagulant/antiplatelet/thrombolytic	

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
J1327	Injection, eptifibatide, 5 mg
J1644	Injection, heparin sodium, per 1000 units
J1645	Injection, dalteparin sodium, per 2500 IU
J1650	Injection, enoxaparin sodium, 10 mg
J3101	Injection, tenecteplase, 1 mg
Drugs administered: anticonvulsant	
J1165	Injection, phenytoin sodium, per 50 mg
J1953	Injection, levetiracetam, 10 mg
Q2009	Injection, fosphenytoin, 50 mg phenytoin equivalent
Drugs administered: antiemetic	
J0780	Injection, prochlorperazine, up to 10 mg
J1790	Injection, droperidol, up to 5 mg
J2405	Injection, ondansetron hydrochloride, per 1 mg
J2550	Injection, promethazine hcl, up to 50 mg
Q0170	Promethazine hydrochloride, 25 mg, oral
Drugs administered: antipsychotic	
J2794	Injection, risperidone, long acting, 0.5 mg
J2950	Injection, promazine hcl, up to 25 mg
J3230	Injection, chlorpromazine hcl, up to 50 mg
Drugs administered: opioid analgesic	
J1170	Injection, hydromorphone, up to 4 mg
J2180	Injection, meperidine and promethazine hcl, up to 50 mg
J2270	Injection, morphine sulfate, up to 10 mg
J2271	Injection, morphine sulfate, 100mg
J2275	Injection, morphine sulfate (preservative-free sterile solution), per 10 mg
J3010	Injection, fentanyl citrate, 0.1 mg
Drugs administered: other drug types	
J0515	Injection, benzotropine mesylate, per 1 mg
J3105	Injection, terbutaline sulfate, up to 1 mg
J0171	Injection, adrenalin, epinephrine, 0.1 mg
J0330	Injection, succinylcholine chloride, up to 20 mg
J0360	Injection, hydralazine hcl, up to 20 mg

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
J1265	Injection, dopamine hcl, 40 mg
J1610	Injection, glucagon hydrochloride, per 1 mg
J1630	Injection, haloperidol, up to 5 mg
J1631	Injection, haloperidol decanoate, per 50 mg
J2250	Injection, midazolam hydrochloride, per 1 mg
J2310	Injection, naloxone hydrochloride, per 1 mg
J2370	Injection, phenylephrine hcl, up to 1 ml
J2710	Injection, neostigmine methylsulfate, up to 0.5 mg
J2765	Injection, metoclopramide hcl, up to 10 mg
J3430	Injection, phytonadione (vitamin k), per 1 mg
S0164	Injection, pantoprazole sodium, 40 mg
Drainage of abscesses and hematomas	
21501	Incision and drainage, deep abscess or hematoma, soft tissues of neck or thorax;
23930	Incision and drainage, upper arm or elbow area; deep abscess or hematoma
25028	Incision and drainage, forearm and/or wrist; deep abscess or hematoma
26010	Drainage of finger abscess; simple
26011	Drainage of finger abscess; complicated (eg, felon)
42700	Incision and drainage abscess; peritonsillar
55100	Drainage of scrotal wall abscess
Laboratory tests	
80100	Drug screen, qualitative; multiple drug classes chromatographic method, each procedure
80196	Salicylate level
82003	Acetaminophen level
82009	Acetone or other ketone bodies, serum; qualitative
82010	Acetone or other ketone bodies, serum; quantitative
82140	Ammonia level
82553	Creatine kinase (CK), (CPK); MB fraction only
82803	Gases, blood, any combination of pH, pCO ₂ , pO ₂ , CO ₂ , HCO ₃ (including calculated O ₂ saturation);
82805	Gases, blood, any combination of pH, pCO ₂ , pO ₂ , CO ₂ , HCO ₃ (including calculated O ₂ saturation); with O ₂ saturation, by direct measurement, except pulse oximetry
83605	Lactate (lactic acid)
83874	Myoglobin
84145	Assay of procalcitonin
84484	Troponin, quantitative
84512	Troponin, qualitative

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
85378	Fibrin degradation products, D-dimer; qualitative or semiquantitative
85379	Fibrin degradation products, D-dimer; quantitative
85380	Fibrin degradation products, D-dimer; ultrasensitive (eg, for evaluation for venous thromboembolism), qualitative or semiquantitative
Removal of foreign body	
42809	Removal of foreign body from pharynx
43215	Esophagoscopy, rigid or flexible; with removal of foreign body
43247	Upper gastrointestinal endoscopy including esophagus, stomach, and either the duodenum and/or jejunum as appropriate; with removal of foreign body
65220	Removal of foreign body, external eye; corneal, without slit lamp
67938	Removal of embedded foreign body, eyelid
Treatment of fractures and dislocations	
21310	Closed treatment of nasal bone fracture without manipulation
21315	Closed treatment of nasal bone fracture; without stabilization
21320	Closed treatment of nasal bone fracture; with stabilization
21400	Closed treatment of fracture of orbit, except blowout; without manipulation
21450	Closed treatment of mandibular fracture; without manipulation
21800	Closed treatment of rib fracture, uncomplicated, each
21820	Closed treatment of sternum fracture
22305	Closed treatment of vertebral process fracture(s)
23505	Closed treatment of clavicular fracture; with manipulation
23540	Closed treatment of acromioclavicular dislocation; without manipulation
23545	Closed treatment of acromioclavicular dislocation; with manipulation
23605	Closed treatment of proximal humeral (surgical or anatomical neck) fracture; with manipulation, with or without skeletal traction
23625	Closed treatment of greater humeral tuberosity fracture; with manipulation
23650	Closed treatment of shoulder dislocation, with manipulation; without anesthesia
23655	Closed treatment of shoulder dislocation, with manipulation; requiring anesthesia
23675	Closed treatment of shoulder dislocation, with surgical or anatomical neck fracture, with manipulation
24500	Closed treatment of humeral shaft fracture; without manipulation
24505	Closed treatment of humeral shaft fracture; with manipulation, with or without skeletal traction
24530	Closed treatment of supracondylar or transcondylar humeral fracture, with or without intercondylar extension; without manipulation
24535	Closed treatment of supracondylar or transcondylar humeral fracture, with or without intercondylar extension; with manipulation, with or without skin or skeletal traction

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
24576	Closed treatment of humeral condylar fracture, medial or lateral; without manipulation
24600	Treatment of closed elbow dislocation; without anesthesia
24605	Treatment of closed elbow dislocation; requiring anesthesia
24620	Closed treatment of Monteggia type of fracture dislocation at elbow (fracture proximal end of ulna with dislocation of radial head), with manipulation
24640	Closed treatment of radial head subluxation in child, nursemaid elbow, with manipulation
24655	Closed treatment of radial head or neck fracture; with manipulation
24670	Closed treatment of ulnar fracture, proximal end (eg, olecranon or coronoid process[es]); without manipulation
25500	Closed treatment of radial shaft fracture; without manipulation
25505	Closed treatment of radial shaft fracture; with manipulation
25520	Closed treatment of radial shaft fracture and closed treatment of dislocation of distal radioulnar joint (Galeazzi fracture/dislocation)
25530	Closed treatment of ulnar shaft fracture; without manipulation
25535	Closed treatment of ulnar shaft fracture; with manipulation
25560	Closed treatment of radial and ulnar shaft fractures; without manipulation
25565	Closed treatment of radial and ulnar shaft fractures; with manipulation
25605	Closed treatment of distal radial fracture (eg, Colles or Smith type) or epiphyseal separation, includes closed treatment of fracture of ulnar styloid, when performed; with manipulation
25622	Closed treatment of carpal scaphoid (navicular) fracture; without manipulation
25630	Closed treatment of carpal bone fracture (excluding carpal scaphoid [navicular]); without manipulation, each bone
25635	Closed treatment of carpal bone fracture (excluding carpal scaphoid [navicular]); with manipulation, each bone
25650	Closed treatment of ulnar styloid fracture
25660	Closed treatment of radiocarpal or intercarpal dislocation, 1 or more bones, with manipulation
25675	Closed treatment of distal radioulnar dislocation with manipulation
25680	Closed treatment of trans-scaphoperilunar type of fracture dislocation, with manipulation
26605	Closed treatment of metacarpal fracture, single; with manipulation, each bone
26641	Closed treatment of carpometacarpal dislocation, thumb, with manipulation
26670	Closed treatment of carpometacarpal dislocation, other than thumb, with manipulation, each joint; without anesthesia
26700	Closed treatment of metacarpophalangeal dislocation, single, with manipulation; without anesthesia
26725	Closed treatment of phalangeal shaft fracture, proximal or middle phalanx, finger or thumb; with manipulation, with or without skin or skeletal traction, each
26740	Closed treatment of articular fracture, involving metacarpophalangeal or interphalangeal joint; without manipulation, each

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
26742	Closed treatment of articular fracture, involving metacarpophalangeal or interphalangeal joint; with manipulation, each
26755	Closed treatment of distal phalangeal fracture, finger or thumb; with manipulation, each
26770	Closed treatment of interphalangeal joint dislocation, single, with manipulation; without anesthesia
26775	Closed treatment of interphalangeal joint dislocation, single, with manipulation; requiring anesthesia
27200	Closed treatment of coccygeal fracture
27220	Closed treatment of acetabulum (hip socket) fracture(s); without manipulation
27230	Closed treatment of femoral fracture, proximal end, neck; without manipulation
27238	Closed treatment of intertrochanteric, peritrochanteric, or subtrochanteric femoral fracture; without manipulation
27246	Closed treatment of greater trochanteric fracture, without manipulation
27250	Closed treatment of hip dislocation, traumatic; without anesthesia
27252	Closed treatment of hip dislocation, traumatic; requiring anesthesia
27256	Treatment of spontaneous hip dislocation (developmental, including congenital or pathological), by abduction, splint or traction; without anesthesia, without manipulation
27265	Closed treatment of post hip arthroplasty dislocation; without anesthesia
27500	Closed treatment of femoral shaft fracture, without manipulation
27501	Closed treatment of supracondylar or transcondylar femoral fracture with or without intercondylar extension, without manipulation
27502	Closed treatment of femoral shaft fracture, with manipulation, with or without skin or skeletal traction
27508	Closed treatment of femoral fracture, distal end, medial or lateral condyle, without manipulation
27510	Closed treatment of femoral fracture, distal end, medial or lateral condyle, with manipulation
27550	Closed treatment of knee dislocation; without anesthesia
27560	Closed treatment of patellar dislocation; without anesthesia
27750	Closed treatment of tibial shaft fracture (with or without fibular fracture); without manipulation
27752	Closed treatment of tibial shaft fracture (with or without fibular fracture); with manipulation, with or without skeletal traction
27760	Closed treatment of medial malleolus fracture; without manipulation
27762	Closed treatment of medial malleolus fracture; with manipulation, with or without skin or skeletal traction
27780	Closed treatment of proximal fibula or shaft fracture; without manipulation
27781	Closed treatment of proximal fibula or shaft fracture; with manipulation
27788	Closed treatment of distal fibular fracture (lateral malleolus); with manipulation
27808	Closed treatment of bimalleolar ankle fracture (eg, lateral and medial malleoli, or lateral and posterior malleoli or medial and posterior malleoli); without manipulation
27810	Closed treatment of bimalleolar ankle fracture (eg, lateral and medial malleoli, or lateral and posterior malleoli or medial and posterior malleoli); with manipulation

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
27816	Closed treatment of trimalleolar ankle fracture; without manipulation
27818	Closed treatment of trimalleolar ankle fracture; with manipulation
27824	Closed treatment of fracture of weight bearing articular portion of distal tibia (eg, pilon or tibial plafond), with or without anesthesia; without manipulation
27825	Closed treatment of fracture of weight bearing articular portion of distal tibia (eg, pilon or tibial plafond), with or without anesthesia; with skeletal traction and/or requiring manipulation
27842	Closed treatment of ankle dislocation; requiring anesthesia, with or without percutaneous skeletal fixation
28400	Closed treatment of calcaneal fracture; without manipulation
28430	Closed treatment of talus fracture; without manipulation
28450	Treatment of tarsal bone fracture (except talus and calcaneus); without manipulation, each
28475	Closed treatment of metatarsal fracture; with manipulation, each
28495	Closed treatment of fracture great toe, phalanx or phalanges; with manipulation
28515	Closed treatment of fracture, phalanx or phalanges, other than great toe; with manipulation, each
28630	Closed treatment of metatarsophalangeal joint dislocation; without anesthesia
28660	Closed treatment of interphalangeal joint dislocation; without anesthesia
29131	Application of finger splint; dynamic
29505	Application of long leg splint (thigh to ankle or toes)
Other diagnostic and therapeutic procedures	
30905	Control nasal hemorrhage, posterior, with posterior nasal packs and/or cautery, any method; initial
30906	Control nasal hemorrhage, posterior, with posterior nasal packs and/or cautery, any method; subsequent
31500	Intubation, endotracheal, emergency procedure
31603	Tracheostomy, emergency procedure; transtracheal
31605	Tracheostomy, emergency procedure; cricothyroid membrane
31720	Catheter aspiration (separate procedure); nasotracheal
32160	Thoracotomy, major; with cardiac massage
32551	Tube thoracostomy, includes water seal (eg, for abscess, hemothorax, empyema), when performed (separate procedure)
33010	Pericardiocentesis; initial
36400	Venipuncture, younger than age 3 years, necessitating physician's skill, not to be used for routine venipuncture; femoral or jugular vein
36430	Transfusion, blood or blood components
36555	Insertion of non-tunneled centrally inserted central venous catheter; younger than 5 years of age
36556	Insertion of non-tunneled centrally inserted central venous catheter; age 5 years or older
36558	Insertion of tunneled centrally inserted central venous catheter, without subcutaneous port or pump; age 5 years or older

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
36569	Insertion of peripherally inserted central venous catheter (PICC), without subcutaneous port or pump; age 5 years or older
36680	Placement of needle for intraosseous infusion
37195	Thrombolysis, cerebral, by intravenous infusion
43752	Naso- or oro-gastric tube placement, requiring physician's skill and fluoroscopic guidance (includes fluoroscopy, image documentation and report)
43753	Gastric intubation and aspiration(s) therapeutic, necessitating physician's skill (eg, for gastrointestinal hemorrhage), including lavage if performed
49082	Abdominal paracentesis (diagnostic or therapeutic); without imaging guidance
54220	Irrigation of corpora cavernosa for priapism
54450	Foreskin manipulation including lysis of preputial adhesions and stretching
62270	Spinal puncture, lumbar, diagnostic
62272	Spinal puncture, therapeutic, for drainage of cerebrospinal fluid (by needle or catheter)
86927	Fresh frozen plasma, thawing, each unit
87040	Culture, bacterial; blood, aerobic, with isolation and presumptive identification of isolates (includes anaerobic culture, if appropriate)
92950	Cardiopulmonary resuscitation (eg, in cardiac arrest)
92953	Temporary transcutaneous pacing
92960	Cardioversion, elective, electrical conversion of arrhythmia; external
92977	Thrombolysis, coronary; by intravenous infusion
93041	Rhythm ECG, 1-3 leads; tracing only without interpretation and report
93042	Rhythm ECG, 1-3 leads; interpretation and report only
99175	Ipecac or similar administration for individual emesis and continued observation until stomach adequately emptied of poison
0042T	Cerebral perfusion analysis using computed tomography with contrast administration, including post-processing of parametric maps with determination of cerebral blood flow, cerebral blood volume, and mean transit time
G0257	Unscheduled or emergency dialysis treatment for an ESRD patient in a hospital
G0390	Trauma response team associated with hospital critical care service
L1830	Knee orthosis, immobilizer, canvas longitudinal, prefabricated
Radiology and related	
70160	Radiologic examination, nasal bones, complete, minimum of 3 views
70450	Computed tomography, head or brain; without contrast material
70487	Computed tomography, maxillofacial area; with contrast material(s)
70496	Computed tomographic angiography, head, with contrast material(s), including noncontrast images, if performed, and image postprocessing

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
70498	Computed tomographic angiography, neck, with contrast material(s), including noncontrast images, if performed, and image postprocessing
71010	Radiologic examination, chest; single view, frontal
71275	Computed tomographic angiography, chest (noncoronary), with contrast material(s), including noncontrast images, if performed, and image postprocessing
72125	Computed tomography, cervical spine; without contrast material
74022	Radiologic examination, abdomen; complete acute abdomen series, including supine, erect, and/or decubitus views, single view chest
78278	Acute gastrointestinal blood loss imaging
78580	Pulmonary perfusion imaging, particulate
78582	Pulmonary ventilation (eg, aerosol or gas) and perfusion imaging
78584	Pulmonary perfusion imaging, particulate, with ventilation; single breath
78585	Pulmonary perfusion imaging, particulate, with ventilation; rebreathing and washout, with or without single breath
78588	Pulmonary perfusion imaging, particulate, with ventilation imaging, aerosol, 1 or multiple projections
78593	Pulmonary ventilation imaging, gaseous, with rebreathing and washout with or without single breath; single projection
A9540	Technetium Tc-99m macroaggregated albumin, diagnostic, per study dose, up to 10
A9558	Xenon Xe-133 gas, diagnostic, per 10 millicuries
Respiratory services	
94002	Ventilation assist and management, initiation of pressure or volume preset ventilators for assisted or controlled breathing; hospital inpatient/observation, initial day
94003	Ventilation assist and management, initiation of pressure or volume preset ventilators for assisted or controlled breathing; hospital inpatient/observation, each subsequent day
94644	Continuous inhalation treatment with aerosol medication for acute airway obstruction; first hour
94645	Continuous inhalation treatment with aerosol medication for acute airway obstruction; each additional hour (List separately in addition to code for primary procedure)
94660	Continuous positive airway pressure ventilation (CPAP), initiation and management
Sedation services	
99143	Moderate sedation services provided by the same physician performing the diagnostic or therapeutic service that the sedation supports, requiring the presence of an independent trained observer to assist in the monitoring of the patient's level of consciousness and physiological status; younger than 5 years of age, first 30 minutes intra-service time

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
99148	Moderate sedation services, provided by a physician other than the health care professional performing the diagnostic or therapeutic service that the sedation supports; younger than 5 years of age, first 30 minutes intra-service time
99149	Moderate sedation services, provided by a physician other than the health care professional performing the diagnostic or therapeutic service that the sedation supports; age 5 years or older, first 30 minutes intra-service time
99150	Moderate sedation services, provided by a physician other than the health care professional performing the diagnostic or therapeutic service that the sedation supports; each additional 15 minutes intra-service time
Wound debridement and repair	
11010	Debridement including removal of foreign material associated with open fracture(s) and/or dislocation(s); skin and subcutaneous tissues
11011	Debridement including removal of foreign material associated with open fracture(s) and/or dislocation(s); skin, subcutaneous tissue, muscle fascia, and muscle
11012	Debridement including removal of foreign material associated with open fracture(s) and/or dislocation(s); skin, subcutaneous tissue, muscle fascia, muscle, and bone
11760	Repair of nail bed
12001	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); 2.5 cm or less
12002	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); 2.6 cm to 7.5 cm
12004	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); 7.6 cm to 12.5 cm
12005	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); 12.6 cm to 20.0 cm
12006	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); 20.1 cm to 30.0 cm
12007	Simple repair of superficial wounds of scalp, neck, axillae, external genitalia, trunk and/or extremities (including hands and feet); over 30.0 cm
12011	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 2.5 cm or less
12013	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 2.6 cm to 5.0 cm
12014	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 5.1 cm to 7.5 cm

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
12015	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 7.6 cm to 12.5 cm
12016	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 12.6 cm to 20.0 cm
12017	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 20.1 cm to 30.0 cm
12018	Simple repair of superficial wounds of face, ears, eyelids, nose, lips and/or mucous membranes; over 30.0 cm
12034	Repair, intermediate, wounds of scalp, axillae, trunk and/or extremities (excluding hands and feet); 7.6 cm to 12.5 cm
12035	Repair, intermediate, wounds of scalp, axillae, trunk and/or extremities (excluding hands and feet); 12.6 cm to 20.0 cm
12036	Repair, intermediate, wounds of scalp, axillae, trunk and/or extremities (excluding hands and feet); 20.1 cm to 30.0 cm
12037	Repair, intermediate, wounds of scalp, axillae, trunk and/or extremities (excluding hands and feet); over 30.0 cm
12041	Repair, intermediate, wounds of neck, hands, feet and/or external genitalia; 2.5 cm or less
12044	Repair, intermediate, wounds of neck, hands, feet and/or external genitalia; 7.6 cm to 12.5 cm
12045	Repair, intermediate, wounds of neck, hands, feet and/or external genitalia; 12.6 cm to 20.0 cm
12046	Repair, intermediate, wounds of neck, hands, feet and/or external genitalia; 20.1 cm to 30.0 cm
12052	Repair, intermediate, wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 2.6 cm to 5.0 cm
12053	Repair, intermediate, wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 5.1 cm to 7.5 cm
12054	Repair, intermediate, wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 7.6 cm to 12.5 cm
12055	Repair, intermediate, wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 12.6 cm to 20.0 cm
12056	Repair, intermediate, wounds of face, ears, eyelids, nose, lips and/or mucous membranes; 20.1 cm to 30.0 cm
13100	Repair, complex, trunk; 1.1 cm to 2.5 cm
13122	Repair, complex, scalp, arms, and/or legs; each additional 5 cm or less
13150	Repair, complex, eyelids, nose, ears and/or lips; 1.0 cm or less
13153	Repair, complex, eyelids, nose, ears and/or lips; each additional 5 cm or less
20103	Exploration of penetrating wound; extremity
40650	Repair lip, full thickness; vermilion only
40652	Repair lip, full thickness; up to half vertical height
40654	Repair lip, full thickness; over 1/2 vertical height, or complex

Table A.1: ED Indicator procedure list

CPT code	Description of procedure
40830	Closure of laceration, vestibule of mouth; 2.5 cm or less
40831	Closure of laceration, vestibule of mouth; over 2.5 cm or complex
41250	Repair of laceration 2.5 cm or less; floor of mouth and/or anterior 2/3 of tongue
41252	Repair of laceration of tongue, floor of mouth, over 2.6 cm or complex
G0168	Wound closure utilizing tissue adhesive(s) only

Appendix B

Original Billings algorithm compared to Ballard

The Billings algorithm gives a very different picture of the severity of the ED visits in this sample compared to the new algorithm.¹ In total, 31.9% of visits were classified by the Billings algorithm as requiring ED care, including 5.3% that were ambulatory care sensitive. Looking only at visits fully classified by the algorithm, 41.8% were classified as requiring ED care. Just under one-quarter of the visits were not classifiable by the emergence ratings: 18.8% were completely unclassified and 4.9% were only classified into one of the special categories.

1. For this analysis, the Billings algorithm was implemented as specified by the developers using the primary diagnosis code for the visit (Billings 2012). Visits that were classified in the main emergence portion of the algorithm that also fall into a special category are presented only in the main emergence classification. That is, visits in the Special category were not otherwise classified by the algorithm. This is different than the implementation of the algorithm using the sample code provided by the NYU group, but does not contradict anything in the original group's publications and writing on the algorithm.

Table B.1: Billings/NYU algorithm: categorization of Medicare sample ED visits

ED visit year	Non-emergent	Emergent: Primary Care Treatable	Billings/NYU algorithm categories				Special category only	Unclassified	Total
			ED Care Needed: preventable/ avoidable	ED Care Needed: not preventable/ avoidable	Total Fully classified				
2011	191,858.3	384,360.1	70,679.7	345,974.9	992,873.0	64,718.0	245,060.0	2,295,524.1	
row pc	8.36%	16.74%	3.08%	15.07%	43.25%	2.82%	10.68%	100.00%	
col pc	48.95%	49.18%	50.05%	49.31%	49.24%	49.64%	49.23%	86.80%	
total pc	7.25%	14.53%	2.67%	13.08%	37.54%	2.45%	9.27%	86.80%	
classified row pc	19.32%	38.71%	7.12%	34.85%	100.00%	N/A	N/A	N/A	
classified total pc	9.52%	19.06%	3.51%	17.16%	49.24%	N/A	N/A	N/A	
2012	200,104.6	397,104.9	70,542.4	355,720.1	1,023,472.0	65,660.0	252,762.0	1,341,894.0	
row pc	14.91%	29.59%	5.26%	26.51%	76.27%	4.89%	18.84%	100.00%	
col pc	51.05%	50.82%	49.95%	50.69%	50.76%	50.36%	50.77%	50.74%	
total pc	7.57%	15.02%	2.67%	13.45%	38.70%	2.48%	9.56%	50.74%	
classified row pc	19.55%	38.80%	6.89%	34.76%	100.00%	N/A	N/A	N/A	
classified total pc	9.92%	19.69%	3.50%	17.64%	50.76%	N/A	N/A	N/A	
Total	391,962.9	781,465.0	141,222.1	701,695.0	2,016,345.1	130,378.0	497,822.0	2,644,545.1	
row pc	14.82%	29.55%	5.34%	26.53%	76.25%	4.93%	18.82%	100.00%	
col pc	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	

Note: pc=percentage

In this table and text that follows, each visit is allowed to fall into multiple categories. An individual visit could be spread over as many as 3 categories, contributing, for example, 0.37 of a visit to the non-emergent total, 0.46 of a visit to the emergent but primary care treatable total, and 0.17 of a visit to the ED care needed/not preventable category.

Table B.2: Comparison of categorization of Medicare sample ED visits: Billings/NYU algorithm and E&M algorithm

E&M category	Billings/NYU category		
	Primary care treatable	ED care needed	Total
Primary care treatable	268,606	98,364	366,970
row pc	73.20%	26.80%	100.00%
col pc	24.32%	12.43%	19.36%
total pc	14.17%	5.19%	19.36%
ED care needed	835,815	692,959	1,528,744
row pc	54.67%	45.33%	100.00%
col pc	75.68%	87.57%	80.64%
total pc	44.09%	36.55%	80.64%
Total	1,104,421	791,323	1,895,744
row pc	58.26%	41.74%	100.00%
col pc	100.00%	100.00%	100.00%

Note: pc=percentage

Includes only visits fully categorized by both algorithms.

There is significant discordance in how the two algorithms classify individual visits. If we collapse both algorithms into two categories—primary care treatable (which includes the non-emergent and emergent but primary care treatable categories) versus “ED care needed” (which includes both emergency care needed/preventable and emergency care needed/not preventable), and look only at visits classified by both algorithms, we see that 54.7% of visits classified as requiring ED care by the new algorithm were classified as “primary care treatable” by the Billings algorithm (versus 58.3% of all visits). See Table B.2. Similarly, of visits classified as requiring ED care by the Billings algorithm, 12.4% were classified as “primary care treatable” by the new algorithm (versus 19.4% of all visits). In total, the algorithms agree on just 50.7% of visits classified by both algorithms, using these binary versions of the measures.

The Ballard algorithm is a version of the NYU algorithm, but yields quite different

results. The two main differences in the algorithm specifications are that 1) the Ballard algorithm collapses the NYU categories into nonemergent, intermediate, and emergent, applying just one category to each diagnosis code, and 2) the Ballard algorithm uses the most severe of all diagnoses attached to a visit rather than just the primary diagnosis code, which has the effect of reducing the number of visits that are unclassified and increasing the likelihood that a visit is placed in the most severe category. Additionally, the Ballard algorithm does not use the ambulatory care sensitive conditions list to categorize visits as potentially preventable/avoidable.

If we collapse the NYU algorithm results into two categories—grouping non-emergent and emergent but primary care treatable and comparing them to the Ballard “nonemergent” category, and grouping the two emergency care needed NYU categories to compare to the Ballard “emergent” category—and exclude the Ballard intermediate visits (formed by visits whose most severe diagnosis rates 50% in the two least severe categories and 50% in the two more severe categories), the Ballard and Billings versions of the algorithm agree on 62% of visits. See Table B.3.

Table B.3: Comparison of categorization of Medicare sample ED visits: condensed Billings/NYU algorithm and Ballard algorithm

Ballard	Original Billings/NYU specification, condensed			
	Primary Care	Treatable	ED Care Needed	Total
Nonemergent	498,905		108,174	607,079
row pc	82.18%		17.82%	100.00%
col pc	44.42%		13.21%	31.25%
total pc	25.68%		5.57%	31.25%
Emergent	624,366		710,989	1,335,355
row pc	46.76%		53.24%	100.00%
col pc	55.58%		86.79%	68.75%
total pc	32.14%		36.60%	68.75%
Total	1,123,271		819,163	1,942,434
row pc	58.20%		41.80%	100.00%
col pc	100.00%		100.00%	100.00%

Note: pc=percentage

Includes only visits fully categorized by both algorithms.

Table B.4: Billings and Ballard versions of the algorithm compared

Ballard version		Billings/NYU version				Total fully classified	Special category only	Unclassified	Total
		Non-emergent	Emergent–primary care treatable	Emerg. care needed: prevent/avoid	Emerg. care needed: not prevent/avoid				
Nonemergent	n	222,444	276,460	22,010	86,164	607,079	32,036	120,678	759,793
	row pc	29.28%	36.39%	2.90%	11.34%	79.90%	4.22%	15.88%	100.00%
	col pc	56.75%	35.38%	15.59%	12.28%	30.11%	24.57%	24.24%	28.73%
	total pc	8.41%	10.45%	0.83%	3.26%	22.96%	1.21%	4.56%	28.73%
Intermediate	n	14,784	35,373	5,323	18,431	73,911	6,018	16,624	96,553
	row pc	15.31%	36.64%	5.51%	19.09%	76.55%	6.23%	17.22%	100.00%
	col pc	3.77%	4.53%	3.77%	2.63%	3.67%	4.62%	3.34%	3.65%
	total pc	0.56%	1.34%	0.20%	0.70%	2.79%	0.23%	0.63%	3.65%
Emergent	n	154,734	469,632	113,889	597,100	1,335,355	76,857	292,046	1,704,258
	row pc	9.08%	27.56%	6.68%	35.04%	35.04%	4.51%	17.14%	100.00%
	col pc	39.48%	60.10%	80.65%	85.09%	85.09%	58.95%	58.66%	64.44%
	total pc	5.85%	17.76%	4.31%	22.58%	22.58%	2.91%	11.04%	64.44%
Total fully classified	n	391,963	781,465	141,222	701,695	2,016,345	N/A	N/A	N/A
	row pc	19.44%	38.76%	7.00%	34.80%	100.00%	N/A	N/A	N/A
	col pc	100.00%	100.00%	100.00%	100.00%	100.00%	N/A	N/A	N/A
	total pc	14.82%	29.55%	5.34%	26.53%	76.25%	N/A	N/A	N/A
Special category only	n	0	0	0	0	0	15,467	4,098	19,565
	row pc	0.00%	0.00%	0.00%	0.00%	0.00%	79.05%	20.95%	100.00%
	col pc	0.00%	0.00%	0.00%	0.00%	0.00%	11.86%	0.82%	0.74%
	total pc	0.00%	0.00%	0.00%	0.00%	0.00%	0.58%	0.15%	0.74%
Unclassified	n	0	0	0	0	0	0	64,376	64,376
	row pc	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	100.00%
	col pc	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.93%	2.43%
	total pc	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.43%	2.43%
Total	n	391,963	781,465	141,222	701,695	2,016,345	130,378	497,822	2,644,545
	row pc	14.82%	29.55%	5.34%	26.53%	76.25%	4.93%	18.82%	100.00%
	col pc	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Note: pc=percentage

A complete tabulation of all Billings/NYU categories vs. the Ballard algorithm categories is available in table B.4. Of note from that table, there is an interesting difference in how the two algorithms rate the less severe categories. 39% of all NYU non-emergent visits are classified as “emergent” by the Ballard algorithm, while 60% of NYU emergent but primary care treatable visits are classified as “emergent” by the Ballard algorithm. (The two NYU emergency care needed categories yield more similar Ballard ratings, both over 80%.) In other words, there is a substantial difference in the Ballard ratings of the two least severe of the original Billings/NYU categories. This difference is driven by the Ballard algorithm’s use of the most severe rating across all diagnoses coded to a visit where the original algorithm uses only the primary diagnosis. The visits coded as emergent but primary care treatable by the NYU algorithm must have more severe diagnoses in the secondary diagnosis fields than those coded as non-emergent by the NYU algorithm. We can’t know from this analysis which of the two algorithms is more accurately describing these ED visits.

Appendix C

Primary care E&M procedures

CPT codes	Visit type
99201-5	Outpatient visit: new patient
99211-5	Outpatient visit: established patient
99304-10	Nursing facility care
99315-18	Nursing home discharge
99324-8	Assisted living/domiciliary care: new patient
99334-7	Assisted living/domiciliary care: established patient
99339-45	Care plan oversight: assisted living/domiciliary/home care
99347-50	Home visits (not hospice or home health agency)
G0402	Welcome to Medicare Preventive Visit
G0438-9	Annual wellness visit

Source: CMS (2012d)

Appendix D

HRR groups

Table D.1: Hospital Referral Region (HRR) groups

Dartmouth HRR	New HRR group	HRR name	HRR group name
1	1001	AL - Birmingham	AL - Birmingham/Huntsville
2	1002	AL - Dothan	AL - Dothan
5	1001	AL - Huntsville	AL - Birmingham/Huntsville
6	1003	AL - Mobile	AL - Mobile
7	1004	AL - Montgomery	AL - Montgomery
9	1005	AL - Tuscaloosa	AL - Tuscaloosa
10	1006	AK - Anchorage	AK - Anchorage
11	1007	AZ - Mesa	AZ - All
12	1007	AZ - Phoenix	AZ - All
14	1007	AZ - Sun City	AZ - All
15	1007	AZ - Tucson	AZ - All
16	1009	AR - Fort Smith	AR - Fort Smith
18	1010	AR - Jonesboro	MO - St. Louis/Columbia/Jonesboro AR/Paducah KY
19	1011	AR - Little Rock	AR - Little Rock
21	1012	AR - Springdale	AR - Springdale
22	1013	AR - Texarkana	AR - Texarkana

Dartmouth HRR	New HRR group	HRR name	HRR group name
23	1014	CA - Orange County	CA - San Diego/Orange Cty/Palm Springs/San Bernardino
25	1015	CA - Bakersfield	CA - Bakersfield
31	1016	CA - Chico	CA - Chico
33	1017	CA - Contra Costa County	CA - San Francisco/Contra Costa Cty
43	1018	CA - Fresno	CA - Fresno
56	1019	CA - Los Angeles	CA - Los Angeles
58	1020	CA - Modesto	CA - Modesto
62	1021	CA - Napa	CA - Napa
65	1022	CA - Alameda County	CA - Alameda County
69	1014	CA - Palm Springs/Rancho Mirage	CA - San Diego/Orange Cty/Palm Springs/San Bernardino
73	1024	CA - Redding	CA - Redding
77	1025	CA - Sacramento	CA - Sacramento/Stockton
78	1026	CA - Salinas	CA - Salinas
79	1014	CA - San Bernardino	CA - San Diego/Orange Cty/Palm Springs/San Bernardino
80	1014	CA - San Diego	CA - San Diego/Orange Cty/Palm Springs/San Bernardino
81	1017	CA - San Francisco	CA - San Francisco/Contra Costa Cty
82	1027	CA - San Jose	CA - San Jose/San Mateo Cty
83	1028	CA - San Luis Obispo	CA - San Luis Obispo
85	1027	CA - San Mateo County	CA - San Jose/San Mateo Cty
86	1030	CA - Santa Barbara	CA - Santa Barbara
87	1031	CA - Santa Cruz	CA - Santa Cruz
89	1032	CA - Santa Rosa	CA - Santa Rosa
91	1025	CA - Stockton	CA - Sacramento/Stockton
96	1034	CA - Ventura	CA - Ventura
101	1035	CO - Boulder	CO - Boulder

Dartmouth HRR	New HRR group	HRR name	HRR group name
102	1036	CO - Colorado Springs	CO - Colorado Springs
103	1037	CO - Denver	CO - Denver/Ft Collins/Greeley
104	1037	CO - Fort Collins	CO - Denver/Ft Collins/Greeley
105	1039	CO - Grand Junction	CO - Grand Junction
106	1037	CO - Greeley	CO - Denver/Ft Collins/Greeley
107	1040	CO - Pueblo	CO - Pueblo
109	1041	CT - Bridgeport	CT - Bridgeport/Hartford/New Haven
110	1041	CT - Hartford	CT - Bridgeport/Hartford/New Haven
111	1041	CT - New Haven	CT - Bridgeport/Hartford/New Haven
112	1043	DE - Wilmington	DE - Wilmington
113	1044	DC - Washington	DC - Washington/Baltimore MD/Takoma Park MD
115	1045	FL - Bradenton	FL - Bradenton
116	1046	FL - Clearwater	FL - Clearwater/St. Petersburg
118	1047	FL - Fort Lauderdale	FL - Ft Lauderdale/Miami
119	1048	FL - Fort Myers	FL - Ft Myers/Sarasota
120	1049	FL - Gainesville	FL - Gainesville
122	1050	FL - Hudson	FL - Tampa/Hudson
123	1051	FL - Jacksonville	FL - Jacksonville
124	1052	FL - Lakeland	FL - Orlando/Lakeland/Ormond Beach
127	1047	FL - Miami	FL - Ft Lauderdale/Miami
129	1053	FL - Ocala	FL - Ocala
130	1052	FL - Orlando	FL - Orlando/Lakeland/Ormond Beach
131	1052	FL - Ormond Beach	FL - Orlando/Lakeland/Ormond Beach
133	1054	FL - Panama City	FL - Panama City
134	1055	FL - Pensacola	FL - Pensacola
137	1048	FL - Sarasota	FL - Ft Myers/Sarasota
139	1046	FL - St. Petersburg	FL - Clearwater/St. Petersburg
140	1057	FL - Tallahassee	FL - Tallahassee
141	1050	FL - Tampa	FL - Tampa/Hudson

Dartmouth HRR	New HRR group	HRR name	HRR group name
142	1058	GA - Albany	GA - Albany
144	1059	GA - Atlanta	GA - Atlanta
145	1060	GA - Augusta	GA - Augusta
146	1061	GA - Columbus	GA - Columbus
147	1062	GA - Macon	GA - Macon
148	1063	GA - Rome	GA - Rome
149	1064	GA - Savannah	GA - Savannah
150	1065	HI - Honolulu	HI - Honolulu
151	1066	ID - Boise	ID - Boise
152	1067	ID - Idaho Falls	ID - Idaho Falls
154	1068	IL - Aurora	IL - Chicago/Evanston/etc./Milwaukee WI
155	1068	IL - Blue Island	IL - Chicago/Evanston/etc./Milwaukee WI
156	1068	IL - Chicago	IL - Chicago/Evanston/etc./Milwaukee WI
158	1068	IL - Elgin	IL - Chicago/Evanston/etc./Milwaukee WI
161	1068	IL - Evanston	IL - Chicago/Evanston/etc./Milwaukee WI
163	1068	IL - Hinsdale	IL - Chicago/Evanston/etc./Milwaukee WI
164	1069	IL - Joliet	IL - Joliet
166	1068	IL - Melrose Park	IL - Chicago/Evanston/etc./Milwaukee WI
170	1070	IL - Peoria	IL - Peoria
171	1068	IL - Rockford	IL - Chicago/Evanston/etc./Milwaukee WI
172	1072	IL - Springfield	IL - Springfield
173	1073	IL - Urbana	IL - Urbana

Dartmouth HRR	New HRR group	HRR name	HRR group name
175	1074	IL - Bloomington	IL - Bloomington
179	1075	IN - Evansville	IN - Evansville
180	1076	IN - Fort Wayne	IN - Fort Wayne
181	1077	IN - Gary	IN - Gary/Munster/South Bend/St. Joseph MI
183	1078	IN - Indianapolis	IN - Indianapolis
184	1079	IN - Lafayette	IN - Lafayette
185	1080	IN - Muncie	IN - Muncie
186	1077	IN - Munster	IN - Gary/Munster/South Bend/St. Joseph MI
187	1077	IN - South Bend	IN - Gary/Munster/South Bend/St. Joseph MI
188	1082	IN - Terre Haute	IN - Terre Haute
190	1083	IA - Cedar Rapids	IA - Cedar Rapids
191	1084	IA - Davenport	IA - Davenport
192	1085	IA - Des Moines	IA - Des Moines
193	1086	IA - Dubuque	IA - Dubuque
194	1087	IA - Iowa City	IA - Iowa City
195	1088	IA - Mason City	IA - Mason City
196	1089	IA - Sioux City	IA - Sioux City
197	1090	IA - Waterloo	IA - Waterloo
200	1091	KS - Topeka	KS - Topeka
201	1092	KS - Wichita	KS - Wichita
203	1093	KY - Covington	KY - Covington
204	1094	KY - Lexington	KY - Lexington
205	1095	KY - Louisville	KY - Louisville
207	1096	KY - Owensboro	KY - Owensboro
208	1010	KY - Paducah	MO - St. Louis/Columbia/Jonesboro AR/Paducah KY
209	1098	LA - Alexandria	LA - Alexandria

Dartmouth HRR	New HRR group	HRR name	HRR group name
210	1099	LA - Baton Rouge	LA - Baton Rouge
212	1100	LA - Houma	LA - Houma
213	1101	LA - Lafayette	LA - Lafayette
214	1102	LA - Lake Charles	LA - Lake Charles
216	1103	LA - Metairie	LA - New Orleans/Metairie/ Slidell
217	1104	LA - Monroe	LA - Monroe
218	1103	LA - New Orleans	LA - New Orleans/Metairie/ Slidell
219	1105	LA - Shreveport	LA - Shreveport
220	1103	LA - Slidell	LA - New Orleans/Metairie/ Slidell
221	1106	ME - Bangor	ME - Bangor
222	1107	ME - Portland	ME - Portland
223	1044	MD - Baltimore	DC - Washington/Baltimore MD/Takoma Park MD
225	1108	MD - Salisbury	MD - Salisbury
226	1044	MD - Takoma Park	DC - Washington/Baltimore MD/Takoma Park MD
227	1109	MA - Boston	MA - Boston/Worcester/Providence RI
230	1110	MA - Springfield	NY - Albany/Binghamton/Springfield MA
231	1109	MA - Worcester	MA - Boston/Worcester/Providence RI
232	1112	MI - Ann Arbor	MI - Detroit/Ann Arbor/Dearborn/Pontiac/Royal Oak
233	1112	MI - Dearborn	MI - Detroit/Ann Arbor/Dearborn/Pontiac/Royal Oak
234	1112	MI - Detroit	MI - Detroit/Ann Arbor/Dearborn/Pontiac/Royal Oak
235	1114	MI - Flint	MI - Flint
236	1115	MI - Grand Rapids	MI - Grand Rapids/Muskegon
238	1116	MI - Kalamazoo	MI - Kalamazoo
239	1117	MI - Lansing	MI - Lansing

Dartmouth HRR	New HRR group	HRR name	HRR group name
240	1118	MI - Marquette	MI - Marquette
242	1115	MI - Muskegon	MI - Grand Rapids/Muskegon
243	1120	MI - Petoskey	MI - Petoskey
244	1112	MI - Pontiac	MI - Detroit/Ann Arbor/Dearborn/Pontiac/Royal Oak
245	1112	MI - Royal Oak	MI - Detroit/Ann Arbor/Dearborn/Pontiac/Royal Oak
246	1121	MI - Saginaw	MI - Saginaw
248	1077	MI - St. Joseph	IN - Gary/Munster/South Bend/St. Joseph MI
249	1122	MI - Traverse City	MI - Traverse City
250	1123	MN - Duluth	MN - Duluth
251	1124	MN - Minneapolis	MN - Mpls/St Paul
253	1125	MN - Rochester	MN - Rochester
254	1126	MN - St. Cloud	MN - St. Cloud
256	1124	MN - St. Paul	MN - Mpls/St Paul
257	1127	MS - Gulfport	MS - Gulfport
258	1128	MS - Hattiesburg	MS - Hattiesburg
259	1129	MS - Jackson	MS - Jackson
260	1130	MS - Meridian	MS - Meridian
261	1131	MS - Oxford	MS - Oxford
262	1132	MS - Tupelo	MS - Tupelo
263	1133	MO - Cape Girardeau	MO - Cape Girardeau
264	1010	MO - Columbia	MO - St. Louis/Columbia/Jonesboro AR/Paducah KY
267	1135	MO - Joplin	MO - Joplin
268	1136	MO - Kansas City	MO - Kansas City
270	1137	MO - Springfield	MO - Springfield
273	1010	MO - St. Louis	MO - St. Louis/Columbia/Jonesboro AR/Paducah KY

Dartmouth HRR	New HRR group	HRR name	HRR group name
274	1139	MT - Billings	MT - Billings
275	1140	MT - Great Falls	MT - Great Falls
276	1141	MT - Missoula	MT - Missoula
277	1142	NE - Lincoln	NE - Lincoln
278	1143	NE - Omaha	NE - Omaha
279	1144	NV - Las Vegas	NV - Las Vegas
280	1145	NV - Reno	NV - Reno
281	1146	NH - Lebanon	NH - Lebanon
282	1147	NH - Manchester	NH - Manchester
283	1148	NJ - Camden	PA - Philadelphia/Newark/White Plains/etc.
284	1148	NJ - Hackensack	PA - Philadelphia/Newark/White Plains/etc.
285	1148	NJ - Morristown	PA - Philadelphia/Newark/White Plains/etc.
288	1148	NJ - New Brunswick	PA - Philadelphia/Newark/White Plains/etc.
289	1148	NJ - Newark	PA - Philadelphia/Newark/White Plains/etc.
291	1148	NJ - Paterson	PA - Philadelphia/Newark/White Plains/etc.
292	1148	NJ - Ridgewood	PA - Philadelphia/Newark/White Plains/etc.
293	1151	NM - Albuquerque	NM - Albuquerque
295	1110	NY - Albany	NY - Albany/Binghamton/Springfield MA
296	1110	NY - Binghamton	NY - Albany/Binghamton/Springfield MA
297	1153	NY - Bronx	NY - Manhattan/Bronx/E Long Island

Dartmouth HRR	New HRR group	HRR name	HRR group name
299	1154	NY - Buffalo	PA - Pittsburgh/Erie/ Buffalo NY/Youngstown OH
300	1155	NY - Elmira	NY - Elmira/Syracuse
301	1153	NY - East Long Island	NY - Manhattan/Bronx/E Long Island
303	1153	NY - Manhattan	NY - Manhattan/Bronx/E Long Island
304	1158	NY - Rochester	NY - Rochester/Sayre PA
307	1155	NY - Syracuse	NY - Elmira/Syracuse
308	1148	NY - White Plains	PA - Philadelphia/Newark/White Plains/etc.
309	1159	NC - Asheville	NC - Asheville
311	1160	NC - Charlotte	NC - Charlotte/Greensboro/Winston-Salem
312	1161	NC - Durham	NC - Raleigh/Durham
313	1160	NC - Greensboro	NC - Charlotte/Greensboro/Winston-Salem
314	1163	NC - Greenville	NC - Greenville
315	1164	NC - Hickory	NC - Hickory
318	1161	NC - Raleigh	NC - Raleigh/Durham
319	1166	NC - Wilmington	NC - Wilmington
320	1160	NC - Winston-Salem	NC - Charlotte/Greensboro/Winston-Salem
321	1167	ND - Bismarck	ND - Bismarck
322	1168	ND - Fargo/Moorhead MN	ND - Fargo/Moorhead MN
323	1169	ND - Grand Forks	ND - Grand Forks
324	1170	ND - Minot	ND - Minot
325	1171	OH - Akron	OH - Cleveland/Akron
326	1172	OH - Canton	OH - Canton
327	1173	OH - Cincinnati	OH - Cincinnati/Dayton/Kettering
328	1171	OH - Cleveland	OH - Cleveland/Akron

Dartmouth HRR	New HRR group	HRR name	HRR group name
329	1175	OH - Columbus	OH - Columbus
330	1173	OH - Dayton	OH - Cincinnati/Dayton/Kettering
331	1176	OH - Elyria	OH - Elyria
332	1173	OH - Kettering	OH - Cincinnati/Dayton/Kettering
334	1177	OH - Toledo	OH - Toledo
335	1154	OH - Youngstown	PA - Pittsburgh/Erie/ Buffalo NY/Youngstown OH
336	1178	OK - Lawton	OK - Lawton
339	1179	OK - Oklahoma City	OK - Oklahoma City
340	1180	OK - Tulsa	OK - Tulsa
341	1181	OR - Bend	OR - Bend
342	1182	OR - Eugene	OR - Eugene
343	1183	OR - Medford	OR - Medford
344	1184	OR - Portland	OR - Portland/Salem
345	1184	OR - Salem	OR - Portland/Salem
346	1186	PA - Allentown	PA - Allentown/Reading/Wilkes-Barre
347	1187	PA - Altoona	PA - Altoona
350	1188	PA - Danville	PA - Danville
351	1154	PA - Erie	PA - Pittsburgh/Erie/ Buffalo NY/Youngstown OH
352	1189	PA - Harrisburg	PA - Harrisburg
354	1190	PA - Johnstown	PA - Johnstown
355	1148	PA - Lancaster	PA - Philadelphia/Newark/White Plains/etc.
356	1148	PA - Philadelphia	PA - Philadelphia/Newark/White Plains/etc.
357	1154	PA - Pittsburgh	PA - Pittsburgh/Erie/ Buffalo NY/Youngstown OH
358	1186	PA - Reading	PA - Allentown/Reading/Wilkes-Barre
359	1158	PA - Sayre	NY - Rochester/Sayre PA

Dartmouth HRR	New HRR group	HRR name	HRR group name
360	1193	PA - Scranton	PA - Scranton
362	1186	PA - Wilkes-Barre	PA - Allentown/Reading/Wilkes-Barre
363	1195	PA - York	PA - York
364	1109	RI - Providence	MA - Boston/Worcester/Providence RI
365	1197	SC - Charleston	SC - Charleston/Columbia/Florence
366	1197	SC - Columbia	SC - Charleston/Columbia/Florence
367	1197	SC - Florence	SC - Charleston/Columbia/Florence
368	1200	SC - Greenville	SC - Greenville
369	1201	SC - Spartanburg	SC - Spartanburg
370	1202	SD - Rapid City	SD - Rapid City
371	1203	SD - Sioux Falls	SD - Sioux Falls
373	1204	TN - Chattanooga	TN - Chattanooga
374	1205	TN - Jackson	TN - Jackson
375	1206	TN - Johnson City	TN - Johnson City
376	1207	TN - Kingsport	TN - Kingsport
377	1208	TN - Knoxville	TN - Knoxville
379	1209	TN - Memphis	TN - Memphis
380	1210	TN - Nashville	TN - Nashville
382	1211	TX - Abilene	TX - Abilene
383	1212	TX - Amarillo	TX - Amarillo
385	1213	TX - Austin	TX - Austin/Temple
386	1214	TX - Beaumont	TX - Beaumont
388	1215	TX - Bryan	TX - Bryan
390	1216	TX - Corpus Christi	TX - Corpus Christi
391	1217	TX - Dallas	TX - Dallas/Ft Worth
393	1218	TX - El Paso	TX - El Paso
394	1217	TX - Fort Worth	TX - Dallas/Ft Worth
396	1220	TX - Harlingen	TX - Harlingen/McAllen
397	1221	TX - Houston	TX - Houston
399	1222	TX - Longview	TX - Longview

Dartmouth HRR	New HRR group	HRR name	HRR group name
400	1223	TX - Lubbock	TX - Lubbock
402	1220	TX - McAllen	TX - Harlingen/McAllen
406	1225	TX - Odessa	TX - Odessa
411	1226	TX - San Angelo	TX - San Angelo
412	1227	TX - San Antonio	TX - San Antonio
413	1213	TX - Temple	TX - Austin/Temple
416	1229	TX - Tyler	TX - Tyler
417	1230	TX - Victoria	TX - Victoria
418	1231	TX - Waco	TX - Waco
420	1232	TX - Wichita Falls	TX - Wichita Falls
421	1233	UT - Ogden	UT - Salt Lake City/Ogden
422	1234	UT - Provo	UT - Provo
423	1233	UT - Salt Lake City	UT - Salt Lake City/Ogden
424	1235	VT - Burlington	VT - Burlington
426	1236	VA - Arlington	VA - Arlington
427	1237	VA - Charlottesville	VA - Charlottesville
428	1238	VA - Lynchburg	VA - Lynchburg
429	1239	VA - Newport News	VA - Newport News
430	1240	VA - Norfolk	VA - Norfolk
431	1241	VA - Richmond	VA - Richmond
432	1242	VA - Roanoke	VA - Roanoke
435	1243	VA - Winchester	VA - Winchester
437	1244	WA - Everett	WA - Seattle/Everett
438	1245	WA - Olympia	WA - Olympia
439	1244	WA - Seattle	WA - Seattle/Everett
440	1246	WA - Spokane	WA - Spokane
441	1247	WA - Tacoma	WA - Tacoma
442	1248	WA - Yakima	WA - Yakima
443	1249	WV - Charleston	WV - Charleston
444	1250	WV - Huntington	WV - Huntington

Dartmouth HRR	New HRR group	HRR name	HRR group name
445	1251	WV - Morgantown	WV - Morgantown
446	1252	WI - Appleton	WI - Appleton
447	1253	WI - Green Bay	WI - Green Bay
448	1254	WI - La Crosse	WI - La Crosse
449	1255	WI - Madison	WI - Madison
450	1256	WI - Marshfield	WI - Marshfield
451	1068	WI - Milwaukee	IL - Chicago/Evanston/etc./Milwaukee WI
452	1257	WI - Neenah	WI - Neenah
456	1258	WI - Wausau	WI - Wausau
457	1259	WY - Casper	WY - Casper

Source: Dartmouth Atlas (2015)

Appendix E

Spatial analysis

To determine whether a spatial analysis is more appropriate than a GEE model for the HRR-level analysis, a spatial error model for panel data was fitted, using the `xsmle` command (Belotti et al. 2013). The model allows errors to be spatially correlated among neighboring geographic units. In particular, the model is as follows:

$$y_{it} = \beta x_{it} + \alpha_i + \gamma_t + \nu_{it} \tag{E.1}$$

$$\nu_{it} = \lambda E\nu_{it} + u_{it} \tag{E.2}$$

where y_{it} is the rate of primary care treatable ED visits per 100 beneficiary-months for HRR group i at time t , x_{it} are HRR group-level covariates, α_i is an HRR group random effect, γ_t is a time effect, E is a spatial matrix indicating which HRR groups are neighbors, u_{it} is a normally distributed error term, and ν_{it} is a spatially autocorrelated error term. The model is fit using maximum likelihood. Table E.1 presents the estimated coefficients.

The spatial parameter λ , which is a coefficient on the matrix of spatially correlated error terms, is small and not statistically significantly different than zero. This suggests that the spatial correlation is negligible, so that our GEE estimates are appropriate. The GEE model gives almost the same effect estimate as the spatial error model (as you would expect, given the non-significant estimate for λ). Recall that the adjusted incidence rate of primary care treatable ED visits was 12.1 when the number of primary care visits per beneficiary was 2, and went down to 6.4 with 6 primary care visits per beneficiary. This suggests an average decrease of 1.425 PCT ED visits for each 1 unit increase in primary care visits per beneficiary. This is very close to the estimate of 1.648 from the spatial error model and is well within its 95% confidence interval.

Table E.1: HRR-level spatial regression results

	PCT ED visits per 100 beneficiary-months
Main equation parameter estimates	
Primary care visit rate	-1.648*** [-2.310,-0.985]
Mortality rate	1.600** [0.424,2.775]
Average age	1.089** [0.378,1.800]
Female (percent)	0.186 [-0.0408,0.412]
White (percent)	0.00877 [-0.0440,0.0615]
Intensity of care index	-0.441 [-2.781,1.900]
Primary care phys per 100k popn	0.0237*** [0.0119,0.0355]
Medicaid eligible (percent)	0.183** [0.0718,0.295]
Comorbidity index (avg HCC score)	-11.95 [-24.27,0.379]
Constant	-79.81*** [-126.7,-32.94]
Spatial parameter estimate	
λ	0.0351 [-0.00129,0.0716]
Random effects variance parameter estimates	
$\ln(\phi)$	1.411*** [1.052,1.770]
σ_e^2	1.144*** [0.769,1.520]
Observations	452
95% confidence intervals in brackets	
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$	

Appendix F

Paper Three full model results

Table F.1: Results

	(1)	(2)	(3)	(4)	(5)	(6)
	PCT ED visits	ED visits	Inpt days	Specialist visits	SNF days	Hospice days
Migration categories						
N only	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]
S only	0.896*** [0.852,0.943]	0.980 [0.957,1.003]	1.090*** [1.049,1.131]	1.138*** [1.124,1.152]	0.896* [0.819,0.980]	1.885*** [1.699,2.091]
Snowbird	1.669*** [1.589,1.753]	1.278*** [1.245,1.312]	0.721*** [0.682,0.762]	1.331*** [1.313,1.350]	0.338*** [0.283,0.404]	0.130*** [0.110,0.155]
Other	2.059*** [1.913,2.216]	1.813*** [1.746,1.883]	1.024 [0.944,1.112]	1.124*** [1.099,1.150]	0.387*** [0.296,0.505]	0.261*** [0.206,0.331]
PC at loc	0.711*** [0.693,0.729]	0.844*** [0.834,0.855]	0.864*** [0.847,0.881]	1.180*** [1.171,1.188]	2.010*** [1.918,2.106]	0.346*** [0.330,0.363]
Migration category X PC at location						
S only X PC at loc	0.960* [0.926,0.996]	0.912*** [0.897,0.928]	0.914*** [0.889,0.939]	0.999 [0.989,1.009]	1.038 [0.964,1.116]	0.730*** [0.684,0.780]
Snowbird X PC at loc	0.590*** [0.559,0.623]	0.702*** [0.683,0.721]	1.185*** [1.122,1.251]	0.938*** [0.924,0.951]	2.047*** [1.721,2.436]	3.105*** [2.673,3.607]
Other X PC at loc	0.495*** [0.454,0.540]	0.545*** [0.523,0.569]	0.939 [0.862,1.024]	1.096*** [1.070,1.123]	2.231*** [1.712,2.907]	2.471*** [2.040,2.992]
State or region						
AZ	0.854*** [0.806,0.906]	1.181*** [1.151,1.211]	0.944** [0.908,0.982]	1.242*** [1.226,1.258]	0.568*** [0.523,0.618]	1.024 [0.911,1.151]
CA	0.724***	1.140***	0.828***	1.355***	0.749***	0.498***

	(1) PCT ED visits	(2) ED visits	(3) Inpt days	(4) Specialist visits	(5) SNF days	(6) Hospice days
CT	[0.689,0.761] 0.727***	[1.114,1.166] 1.002	[0.801,0.856] 0.935***	[1.340,1.370] 1.419***	[0.702,0.799] 0.900***	[0.444,0.558] 0.575***
FL	[0.692,0.763] 0.596***	[0.979,1.026] 0.935***	[0.902,0.971] 0.876***	[1.401,1.437] 1.462***	[0.853,0.950] 0.805***	[0.505,0.655] 0.779***
GA	[0.570,0.623] 0.773***	[0.916,0.955] 1.149***	[0.849,0.903] 0.911***	[1.447,1.477] 1.200***	[0.758,0.856] 0.704***	[0.707,0.859] 1.102
MA	[0.731,0.818] 0.785***	[1.120,1.179] 1.030**	[0.879,0.946] 1.001	[1.185,1.215] 1.166***	[0.653,0.760] 0.924***	[0.969,1.254] 0.766***
MI	[0.755,0.817] 0.709***	[1.010,1.049] 1.045***	[0.974,1.029] 0.952***	[1.154,1.179] 1.060***	[0.882,0.968] 0.771***	[0.694,0.845] 0.935
MN	[0.685,0.734] 1.175***	[1.028,1.062] 1.067***	[0.930,0.974] 1.008	[1.050,1.071] 0.872***	[0.739,0.805] 1.026	[0.861,1.015] 0.862*
NJ	[1.116,1.237] 0.480***	[1.040,1.094] 0.880***	[0.973,1.044] 0.935***	[0.858,0.887] 1.626***	[0.959,1.097] 0.803***	[0.761,0.977] 0.634***
NY	[0.456,0.505] 0.613***	[0.864,0.895] 0.841***	[0.913,0.958] 0.994	[1.611,1.641] 1.551***	[0.771,0.837] 0.687***	[0.579,0.694] 0.410***
PA	[0.592,0.634] 0.681***	[0.829,0.854] 0.957***	[0.973,1.015] 1.011	[1.538,1.564] 1.129***	[0.661,0.714] 0.805***	[0.375,0.447] 0.901**
SC	[0.658,0.705] 0.839***	[0.942,0.973] 1.049***	[0.989,1.034] 0.985	[1.119,1.140] 1.146***	[0.773,0.838] 0.813***	[0.832,0.975] 1.125
TX	[0.791,0.890] 0.786***	[1.021,1.078] 1.090***	[0.947,1.024] 1.080***	[1.131,1.162] 1.133***	[0.750,0.881] 0.846***	[0.990,1.278] 0.911
	[0.750,0.825]	[1.066,1.114]	[1.045,1.116]	[1.121,1.146]	[0.792,0.903]	[0.818,1.015]

	(1)	(2)	(3)	(4)	(5)	(6)
	PCT ED visits	ED visits	Inpt days	Specialist visits	SNF days	Hospice days
WI	0.881*** [0.837,0.927]	1.062*** [1.039,1.085]	0.901*** [0.875,0.928]	0.962*** [0.950,0.973]	0.965 [0.913,1.020]	1.199*** [1.086,1.324]
Other New England	1.202*** [1.153,1.254]	1.135*** [1.112,1.158]	0.883*** [0.854,0.914]	0.963*** [0.951,0.975]	0.906*** [0.857,0.957]	0.818*** [0.735,0.911]
Other Mideast	0.647*** [0.622,0.674]	0.983 [0.964,1.002]	0.908*** [0.884,0.932]	1.411*** [1.397,1.425]	0.800*** [0.763,0.839]	0.830*** [0.752,0.915]
Other Great Lakes	0.718*** [0.697,0.740]	1.032*** [1.019,1.045]	1.004 [0.987,1.021]	1.193*** [1.184,1.201]	1.049** [1.016,1.084]	0.956 [0.896,1.019]
Other Plains	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]
Other Southeast	0.874*** [0.836,0.913]	1.188*** [1.163,1.212]	1.073*** [1.041,1.107]	1.046*** [1.035,1.057]	0.913** [0.858,0.972]	0.835*** [0.754,0.924]
Other Southwest	1.298*** [1.230,1.369]	1.260*** [1.228,1.293]	1.093*** [1.053,1.136]	0.911*** [0.898,0.924]	0.767*** [0.708,0.831]	1.031 [0.908,1.170]
Rocky Mountain	0.985 [0.947,1.025]	1.120*** [1.099,1.140]	0.920*** [0.896,0.945]	0.996 [0.986,1.007]	0.975 [0.927,1.026]	1.250*** [1.146,1.363]
Far West-N	0.805*** [0.775,0.836]	1.051*** [1.033,1.070]	0.832*** [0.811,0.854]	1.007 [0.997,1.017]	0.769*** [0.731,0.809]	0.944 [0.861,1.035]
Other Far West-S	0.680*** [0.630,0.733]	1.197*** [1.159,1.237]	1.067* [1.010,1.126]	1.197*** [1.176,1.218]	0.693*** [0.627,0.767]	0.909 [0.776,1.065]
Age	0.994*** [0.993,0.994]	1.000* [1.000,1.001]	0.989*** [0.988,0.989]	0.982*** [0.982,0.982]	1.044*** [1.043,1.044]	1.106*** [1.104,1.108]
Season						

	(1)	(2)	(3)	(4)	(5)	(6)
	PCT ED visits	ED visits	Inpt days	Specialist visits	SNF days	Hospice days
Jan-Mar 2011	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]
Apr-Jun 2011	1.062*** [1.044,1.081]	0.996 [0.989,1.003]	0.935*** [0.925,0.945]	1.020*** [1.018,1.023]	0.914*** [0.898,0.931]	1.173*** [1.156,1.190]
Jul-Sep 2011	1.128*** [1.109,1.148]	1.020*** [1.013,1.027]	0.926*** [0.916,0.937]	1.002 [1.000,1.005]	0.931*** [0.913,0.951]	1.415*** [1.390,1.440]
Oct-Dec 2011	0.979* [0.962,0.997]	0.995 [0.988,1.002]	0.938*** [0.927,0.948]	0.954*** [0.951,0.956]	0.939*** [0.920,0.958]	1.629*** [1.598,1.660]
Jan-Mar 2012	0.977* [0.959,0.994]	1.077*** [1.070,1.085]	1.064*** [1.052,1.076]	1.010*** [1.008,1.013]	1.039*** [1.018,1.060]	1.915*** [1.878,1.953]
Apr-Jun 2012	1.070*** [1.051,1.090]	1.091*** [1.083,1.098]	1.013* [1.001,1.025]	1.025*** [1.022,1.027]	1.071*** [1.049,1.093]	2.190*** [2.147,2.233]
Jul-Sep 2012	1.123*** [1.102,1.143]	1.140*** [1.132,1.148]	1.032*** [1.020,1.044]	1.006*** [1.004,1.009]	1.129*** [1.106,1.153]	2.506*** [2.457,2.556]
Oct-Dec 2012	1.008 [0.990,1.027]	1.135*** [1.127,1.143]	1.032*** [1.020,1.044]	0.957*** [0.954,0.960]	1.197*** [1.173,1.222]	2.692*** [2.638,2.746]
White	0.719*** [0.704,0.735]	0.830*** [0.823,0.837]	0.893*** [0.882,0.904]	1.114*** [1.109,1.119]	1.060*** [1.037,1.084]	1.307*** [1.245,1.372]
Medicaid	1.516*** [1.489,1.544]	1.150*** [1.141,1.159]	0.868*** [0.858,0.877]	0.706*** [0.703,0.710]	1.494*** [1.468,1.521]	0.962 [0.921,1.004]
Female	0.943*** [0.931,0.956]	1.037*** [1.031,1.042]	1.022*** [1.014,1.029]	0.948*** [0.945,0.950]	1.295*** [1.276,1.314]	1.187*** [1.153,1.223]
Comorbidities (HCC score quintiles)						

	(1)	(2)	(3)	(4)	(5)	(6)
	PCT ED visits	ED visits	Inpt days	Specialist visits	SNF days	Hospice days
Lowest scores	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]	1 [1,1]
2	1.238*** [1.213,1.265]	1.490*** [1.473,1.507]	2.435*** [2.379,2.492]	1.370*** [1.364,1.377]	2.415*** [2.241,2.602]	2.233*** [2.026,2.461]
3	1.401*** [1.372,1.431]	2.070*** [2.047,2.094]	5.053*** [4.945,5.163]	1.687*** [1.679,1.695]	6.088*** [5.687,6.517]	3.381*** [3.079,3.713]
4	1.651*** [1.613,1.689]	3.146*** [3.112,3.181]	11.81*** [11.57,12.06]	2.052*** [2.042,2.062]	16.55*** [15.50,17.67]	5.701*** [5.198,6.253]
Highest scores	2.312*** [2.262,2.363]	7.311*** [7.234,7.389]	55.16*** [54.08,56.27]	2.652*** [2.639,2.665]	72.98*** [68.44,77.82]	13.18*** [12.03,14.44]
Constant	0.0296*** [0.0277,0.0317]	0.0346*** [0.0336,0.0356]	0.0581*** [0.0556,0.0607]	0.835*** [0.822,0.849]	2.26E-4*** [2.05E-4,2.50E-4]	2.60E-5*** [2.19E-5,3.10E-5]
Observations	8,892,882	8,892,882	8,892,882	8,892,882	8,892,882	8,892,882

Exponentiated coefficients; 95% confidence intervals in brackets

X denotes interaction terms

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$