Journal of Health Economics 108 (2026) 103149

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Tele-triage, care substitution, and health: Evidence from quasi-randomly assigned nurses[☆]

Liam Rose[a], Ken Suzuki[a,*], Linda Diem Tran[a], Anita Vashi[b]

[a] Department of Veterans Affairs and Stanford Medicine, United States of America
[b] Department of Veterans Affairs, University of California, San Francisco, and Stanford Medicine, United States of America

ARTICLE INFO

Keywords:
Triage
Examiners design
Judges design
Acute care
Patient choices

ABSTRACT

Patients experiencing acute health symptoms often face uncertainty about how and where to receive care. We study patients who call a nurse advice line and receive one of four recommendations: emergency department (ED), urgent care (UC), primary care (PC), or self-care (Home). Leveraging an extension of examiner designs that recovers margin-specific effects for each pair of adjacent recommendations (ED-UC, UC-PC, PC-Home), we estimate the impact of nurse recommendations on both patient decisions and their subsequent health outcomes. We find that recommendations have large impacts on patient decisions at each margin. We then show that UC recommendations reduce 28-day healthcare costs by $404 relative to ED recommendations and by $247 relative to PC recommendations, suggesting substantial potential for cost savings through improved triage.

1. Introduction

Individuals experiencing acute health symptoms often face a difficult decision about how and where to seek care, yet most patients will not have sufficient medical knowledge to know if their symptoms require immediate attention. Some patients may be fearful of high costs often associated with emergency care and avoid seeking immediate attention, while others that present to the emergency department may be just as well off going to urgent care or waiting to see a primary care physician at much lower costs per visit. While ED care is clearly beneficial to patients with the most acute symptoms, much of the care provided in EDs has been identified as a misallocation of resources and a potential source of reducing health care spending.

Telephone triage services are one of the primary ways providers and insurers have sought to aid patients with this decision. These services are increasingly prevalent across healthcare systems and insurance networks, including all major U.S. insurers, national health systems like the NHS 111 and Germany's Der Patientenservice 116 117, and state and federal health agencies. These lines provide individuals with access to medical personnel – most often a registered nurse – who can offer health-related information, advice, and guidance. For patients seeking unplanned care, these services are intended to serve as the first point of contact, where nurses assess symptoms and recommend an appropriate level of care. Beyond offering guidance, telephone triage has become a critical tool for managing rising healthcare demands and alleviating ED overcrowding by directing patients to the most suitable care setting.

☆ We are grateful to Todd Wagner, Carlos Dobkin, Laura Giuliano, Jasmin Moshfegh, Hannah Bae, Nathaniel Breg, and seminar participants at UC Santa Cruz and Stanford University for helpful comments. This project was funded by the Department of Veterans Affairs (Office of Integrated Veteran Care and project EBP 22-108, 1I50HX003614-01). The content is solely the responsibility of the authors. All errors are our own. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

* Corresponding author.
E-mail addresses: liamrose@stanford.edu (L. Rose), ksuzuki4@stanford.edu (K. Suzuki), diemtran@stanford.edu (L.D. Tran), anita.vashi@va.gov (A. Vashi).

https://doi.org/10.1016/j.jhealeco.2026.103149
Received 28 November 2025; Received in revised form 13 April 2026; Accepted 18 May 2026
Available online 22 May 2026

Although telephone triage services can play a potentially important role in reducing information frictions and improving allocative efficiency, we know little about whether and how triage recommendations influence patient choices, healthcare costs, and health outcomes within healthcare delivery. Furthermore, there is little empirical evidence on the consequences of human variability in the upstream stage of access to care. Prior research highlights substantial variation in clinical decisions among healthcare professionals, suggesting that some patients may receive a higher or lower level of care than their symptoms warrant (Abualenain et al., 2013; Molitor, 2018; Chan and Gruber, 2020; Chan et al., 2022; Coussens and Ly, 2025). In the triage context, some nurses may perform triage too defensively, funneling low-acuity patients into the ED unnecessarily. Other nurses may advise patients to seek less intensive care too optimistically, causing a delay in timely access to critical care.

We aim to address this gap using unique administrative data from a nurse advice line operated by the Veterans Affairs (VA) healthcare system. We apply an examiners design, leveraging the quasi-random assignment of calls to the next available nurse within call centers. In this context, triage nurses choose one of four follow-up recommendations: emergency department (ED), urgent care (UC), primary care (PC), or self-care (Home) according to symptom. Unlike most U.S. settings, patients face low or no cost sharing that is equal in each possible care setting.

These margins are particularly important for several reasons. First, ED care accounts for over 5% of U.S. healthcare spending, and it has been estimated that more than 30% of ED visits could be managed in less acute settings (Weinick et al., 2010; Vashi et al., 2019). This has led to substantial efforts to reduce this type of misallocation, including the proliferation of nurse triage lines.[1] Second, urgent care centers have been shown to substitute for EDs, particularly when examined through availability during extended hours (Allen et al., 2021) or when clinics become available in a local market (Alexander et al., 2019). However, UCs may also contribute to higher overall healthcare spending through downstream hospital admissions (Currie et al., 2023). Nevertheless, urgent care centers have proliferated quickly, nearly doubling in number in the US between 2013 and 2023. This may also be driven by increased wait times for primary care appointments, highlighting the importance of urgent care centers for those on the lower-acuity margin as well (Morgan et al., 2026). However, unlike previous works, patients on either margin do not have a financial incentive to choose one care setting over another.

However, this institutional setting presents two challenges vis-a-vis the traditional examiner design. First, although the conventional examiners design literature often collapses multiple treatment alternatives into a binary treatment, recent research highlights that two-stage least squares (TSLS) may fail to identify interpretable causal effects when examiners affect potential treatments on sub-treatment margins (Mueller-Smith, 2015; Chyn et al., 2024). In our context, one might collapse triage recommendations into ED vs. non-ED (UC, PC, and Home) and use nurse ED recommendation propensity as a single instrument for the binarized ED recommendation indicator. However, if nurses with a high ED propensity also tend to recommend UC over PC, for example, this TSLS design does not recover interpretable effects of ED recommendation.[2]

To address this identification problem, we examine identification conditions under which TSLS recovers margin-specific causal effects on each of two adjacent recommendations (ED-UC, UC-PC, PC-Home), based on recent methodological development in the examiners design (Humphries et al., 2024). We use TSLS with a focal recommendation propensity as an instrument for the focal recommendation (e.g., ED), while controlling for non-focal recommendation propensities (e.g., PC and Home). Under the identification assumptions, this method allows us to recover the margin-specific local average treatment effect (LATE) on each margin (e.g., the effect of ED recommendation relative to UC recommendation).

Second, our setting differs from the conventional examiner setup in that patients need not follow the recommendation of the examiner. This is particularly problematic in the multiple treatments setting. Intuitively, assignment to a nurse with a high propensity to recommend ED may affect downstream health outcomes through patient choices other than whether to visit the ED. Being assigned a "high-ED" nurse may influence the probability of seeking primary care or urgent care in addition to the probability of visiting the ED as expected. Changes in downstream outcomes then become the product of multiple channels of behavior moved by the instrument. Because of this, our TSLS estimates instead recover the effects of receiving a particular recommendation (e.g., ED) – and not the effects of receiving a particular type of care.[3]

We find substantial variation in triage recommendations among nurses on the high and middle acuity margins (ED-UC and UC-PC), after controlling for non-focal recommendation propensities and patient and call baseline characteristics. Our first-stage coefficient indicates that reassigning a call from a nurse at the 5th percentile to one at the 95th percentile of the ED recommendation propensity distribution increases the likelihood of being recommended ED over UC by 12.6 percentage points. Similarly, call reassignment from the 5th to the 95th of the UC recommendation propensity distribution increases the likelihood of being recommended UC over PC by 19.3 percentage points. On the low acuity margin (PC-Home), calls assigned to the 5th and the 95th percentile nurses differ by 4.9 percentage points in their propensity to be recommended PC over Home.

We then show that, on average, patients recommended ED care are 24.0 percentage points less likely to use UC and 25.3 percentage points more likely to use ED within three days of the call, compared to those recommended UC. In addition, the ED

[1] Evaluations of other efforts to move care out of EDs aside from urgent care have shown mixed results (Flores-Mateo et al., 2012; Raven et al., 2016). Increases in primary care access successfully reduced ED visits in uninsured and Medicaid populations (Sadowski et al., 2009; Retchin et al., 2009), but case management, individualized care plans, and information sharing were not consistently effective (Soril et al., 2015).

[2] We later show that the TSLS identifies the sum of (i) a weighted average of sub-LATEs involving ED recommendation and (ii) bias terms if we binarize recommendations into ED and non-ED.

[3] Recent studies highlight similar identification problems due to subjects' non-compliance in the context of randomized field experiments with multiple treatment alternatives (Kirkeboen et al., 2016; Kline and Walters, 2016; Pinto, 2022). The literature seeks natural ways to impose identification restrictions on non-compliance. For instance, Kline and Walters (2016) and Pinto (2022) restrict certain response types, invoking revealed preferences. Exploring such restrictions in examiners design is beyond the scope of this paper, and we leave it for future research.

recommendation modestly shifts patient utilization on the extensive margin. Patients recommended ED are 6.9 percent more likely to shift their choice from self-care (Home) to any professional care (PC, UC, or ED), compared to patients recommended UC. These patients are less likely to need to call the triage line again. However, 28 days after the initial call, patients recommended ED care incur $404 greater cumulative health care spending than patients recommended UC.

For patients on the less acute UC-PC margin, we find that being recommended UC care leads to a 11.5 percentage point (32.7%) drop in the primary care visits and a 20.7 percentage point increase in urgent care use within three days of the call. Furthermore, patients recommended UC are 1.2 percentage points (13.8%) less likely to use ED within three days, compared to patients recommended PC. On the extensive margin, the UC recommendation modestly increases the probability of receiving any care by 4.8 percentage points relative to the PC recommendation. Overall, this mixture of shifts in utilization contributes to a reduction in health care spending. After 28 days, patients recommended UC have $247 less in costs, compared to patients recommended PC.

Lastly, on the least acute PC-Home margin, our results suggest that patient utilization shifts almost exclusively on the corresponding care margin. Patients recommended PC are 7.6 percentage points more likely to use primary care within three days, compared to patients recommended Home. By contrast, the effects of being recommended PC over Home on other utilization indicators (UC, ED, and hospital admission) are small and statistically insignificant.

This paper contributes to several strands of literature. First, our study is closely related to the literature on variation in clinical decisions among healthcare professionals (Abualenain et al., 2013; Molitor, 2018; Chan and Gruber, 2020; Coussens and Ly, 2025) and research that exploits quasi-random assignment of healthcare providers and variations in their practice style (Chan et al., 2022; Silver and Zhang, 2022; Chan et al., 2023) or prescribing tendencies (Bhalotra et al., 2025; Currie and Zwiers, 2025; Costa-Ramon et al., 2023; Eichmeyer and Zhang, 2022, 2023; Dalsgaard et al., 2014). Although healthcare providers often face multiple treatment alternatives, applications of the examiners design generally focus on a binary (or binarized) treatment. The assumption that comes with this approach is that instruments induce compliers to take up a specific treatment without inducing changes in other treatments. When substitutes are more readily available the instrument can be correlated with other treatment pathways, such as SSRI propensity with other drug propensity (Bhalotra et al., 2025) or opiate prescribing with non-opioid alternatives (Eichmeyer and Zhang, 2023). While in these cases it may be reasonable to assume that the instruments influence the take-up of the focal treatment and do not meaningfully impact other treatments, the setting presented in this paper is a clear violation of that assumption, such as the ED recommendation from a nurse inducing some patients to visit the ED while also inducing a patient that might have otherwise stayed home to instead visit urgent care or contact their primary care provider. Therefore, we extend recent methodological developments in the judges design into a healthcare setting to address these multiple treatments possibilities (Arteaga, 2023; Humphries et al., 2024; Kamat et al., 2024). Our application is most closely related to Humphries et al. (2024) that considers identification conditions to recover margin-specific causal effects (e.g. incarceration vs. conviction and conviction vs. dismissal) when judges make an ordered choice based on a single latent index.

Second, we contribute to a limited but growing literature on the impact of triage decisions on access to care within healthcare delivery (Chan and Gruber, 2020; Ferro and Serra, 2025; Islam et al., 2021; Sexton et al., 2022). The literature on triage in EDs documents substantial variation in triage decisions between and within triage nurses, with patients assigned a lower priority experiencing longer wait time or additional ED care afterward (Chan and Gruber, 2020; Ferro and Serra, 2025). We show that the telephone triage process and the variability among triage nurses affect patient choices and access to care further upstream, even before patients go to healthcare facilities.

Finally, a broad literature investigates the consequences of substituting care settings. While ED care accounts for over 5% of U.S. healthcare spending, studies suggest that more than 30% of ED visits could be managed in less acute settings (Weinick et al., 2010; Vashi et al., 2019).[4] Urgent care centers have been shown to substitute for EDs, particularly when examined through availability during extended hours (Allen et al., 2021) or when clinics become available in a local market (Alexander et al., 2019). However, UCs may also contribute to higher overall healthcare spending (Wang et al., 2021; Currie et al., 2023). Currie et al. (2023) shows downstream hospital admissions as a potential cause of increased spending. Using a different source of variation, we show that for patients considered by a triage nurse to be on the margin of needing ED care, substitution toward UC can be cost-saving.

2. Background

The VA operates one of the nation's largest health care systems, providing care to approximately 10 million veterans at 171 medical centers and 1113 outpatient facilities distributed across the country. To receive VA health care, an individual must have served and been honorably discharged from the military and qualify under at least one of three broad categories: have a disability connected to their service, have income below a set threshold, or have been discharged within the last five years.[5] In a given year, VA provides care to about one-third of US veterans, providing extensive service in a vertically-integrated system that includes primary care, mental health care, specialty care, acute care, and long-term care.

VA medical centers and outpatient clinics generally operate on a "hub-and-spoke" model, where regional medical centers work together with a number of nearby outpatient clinics. Medical centers are then geographically divided into 18 regional care systems

[4] Evaluations of efforts to move care out of EDs aside from urgent care have shown mixed results (Flores-Mateo et al., 2012; Raven et al., 2016). Increases in primary care access successfully reduced ED visits in uninsured and Medicaid populations (Sadowski et al., 2009; Retchin et al., 2009), but case management, individualized care plans, and information sharing were not consistently effective (Soril et al., 2015).

[5] VA uses the Department of Housing and Urban Development's annual geographic-based income limits, further allowing individuals to be 10% over the threshold if they agree to pay copays. Over 80% of enrolled veterans face no cost sharing.

known as Veteran Integrated Service Networks (VISNs). Historically, medical centers and VISNs developed their own call centers to serve as entry points for veterans and their families. The call centers provide frequently used administrative and clinical services. While services have differed somewhat among call centers, they all provided some form of assistance with appointment scheduling, enrollment questions, pharmacy services, and nurse triage.

Nurse triage services allow patients to speak with a Registered Nurse (RN) for symptom evaluation and healthcare disposition. When a veteran calls for triage, the next available nurse is assigned to assess the patient's needs and recommend appropriate follow-up care. The triage process is standardized through a decision-support algorithm used across all nurses and call centers. First, the nurse gathers and inputs basic patient information – such as age, gender, chief complaint, and pain scale – into the algorithm. Second, the algorithm generates clinical questions based on these inputs. Third, the nurse communicates with the patient and enters the responses. Finally, the algorithm provides recommendations for disposition (e.g., ED, urgent care, primary care, dentist, or self-care) and follow-up timing (e.g., now-911, now, 2–8 h, 12–24 h, 2–3 days, 1–2 weeks). Appendix Figure B1 illustrates these steps.

While the decision-support algorithm standardizes the telephone triage process, nurses still exercise discretion by overriding the algorithm's triage recommendations. This study exploits cross-nurse variations in their discretion within algorithm recommendations. Appendix Figure B2 illustrates this by plotting each nurse's propensity to recommend ED care, conditional on whether the algorithm suggests ED care (y-axis) or non-ED care (x-axis), across call centers. Nurses positioned in the top-right corner of each panel (call center) consistently follow the algorithm, whereas those farther from this corner are more likely to override the algorithmic recommendations.

In addition to providing services directly, VA also purchases care from non-VA providers. Importantly, this includes emergency care, with more than one-third of ED visits involving VA occurring at non-VA facilities. VA encourages enrollees that consider their life or health to be in danger to seek immediate medical attention, and prior approval is not required. Further, VA maintains a network of non-VA urgent care centers that enrollees can utilize. While unable to make diagnoses or recommend specific treatments on these calls, triage nurses are instructed to work with patients to direct them to the appropriate care location, including taking into account the patient's preferences. VA advice on when to use these two services aligns with that of other providers and insurers, namely that urgent care can handle medical problems that need to be treated right away but are not immediate threats to health or life. Some examples include sprains, minor broken bones, moderate flu-like symptoms, or mild nausea or vomiting.[6]

3. Data

3.1. Overview

We construct our analysis sample by linking multiple sources of administrative data from the VA, including records of nurse triage cases, healthcare utilization, and patient demographics. This section sketches the most relevant information about our analysis sample. Appendix Table B1 describes our data cleaning and sample construction in further detail.

3.2. Data sources and sample construction

Our sample construction starts with the universe of telephone triage cases received in all call centers across the US from July 1, 2018, to December 31, 2022. The triage records have information at the call level, including triage date-time (year, month, day, hour, and minute), patient ID, triage nurse ID, station (call center) ID, triage disposition (recommended follow-up location and timing), chief complaint (symptom), pain scale (0–10), and duration of chief complaint. Each call is further linked to the patient's prior healthcare utilization events at VA facilities (within 365 days of the triage), prior diagnoses (31 Elixhauser comorbidity indices), VA benefits eligibility status (priority group indicators), and demographics (e.g., age, gender, marital status). We use these covariates for randomization and robustness checks as well as for profiling complier characteristics.

We define a patient is recommended "ED" by nurse if nurse follow-up location is emergency department. Likewise, we say that a patient is recommended "UC" if nurse follow-up location is urgent care. We define a patient to be recommended "PC" if nurse's follow-up location is clinic or other miscellaneous categories, such as virtual care and dentist. Lastly, we say that a patient is recommended "Home" if nurse's follow-up location is home. Appendix Table B2 describes triage recommendations in detail.

Our primary outcomes include healthcare utilization events after triage (primary care, urgent care, and ED), emergency hospital admission (those that came through the ED), repeat calls to the triage line, health care spending, and mortality. We construct mortality indicators from the date of the triage call using VA vital status data, which are further supplemented with death records from both Medicare and the Social Security Administration. We also capture full costs of downstream care. Costs for VA-provided care come from Managerial Cost Accounting (MCA) data, which provides the official cost estimates for VA encounters. Costs for VA-paid care come from claims paid by VA to outside providers.

To construct our main sample, we impose the following key restrictions (See Appendix Table B1 for details). First, we drop the calls during non-business hours (before 8 am, after 4 pm, weekends, and holidays). Some call centers do not offer telephone triage

[6] VA maintains a website to help patients differentiate between urgent and emergent care (https://www.va.gov/resources/choosing-between-urgent-and-emergency-care/), which is similar to other provider's websites on the same topic (e.g. Yale Medicine: https://www.yalemedicine.org/news/should-you-go-to-the-emergency-department-or-urgent-care).

Table 1
Characteristics of baseline sample.

Variable	Mean
Age	63.054
Male	0.863
Married	0.513
White	0.691
Black	0.182
Hispanic	0.061
Rural county	0.206
Comorbidity count	2.901
ED visit in prior year	0.382
Admission in prior year	0.102
Inpatient in prior year	0.101
Primary care in prior year	0.974
Primary care visit 3d	0.307
Urgent care visit 3d	0.063
ED visit 3d	0.192
Admission 3d	0.027
Anycare 3d	0.499
Inpatient cost 3d	153.232
Outpatient cost 3d	286.668
Totalcost 3d	439.900
Nurse recommended PC	0.627
Nurse recommended UC	0.044
Nurse recommended ED	0.277
Algorithm recommended PC	0.667
Algorithm recommended UC	0.004
Algorithm recommended ED	0.263
Calls	1,273,843
Patients	836,420
Nurses	1725
Call centers	96

Notes: This table presents characteristics of calls in the baseline sample.

during non-business hours and transfer calls to other call centers or non-VA contractors. Second, we remove calls from patients with the most recent prior triage call or ED visit within 30 days to define an index triage incident. Third, we only keep calls in nurse-by-call center-by-year-by-algorithm disposition cells with at least 10 observations.[7] From this we drop any cells that do not have at least two nurses remaining. With these restrictions, our baseline sample consists of 1,273,843 calls (from 836,420 patients) received by 1725 nurses at 96 call centers.

Table 1 summarizes the characteristics of our sample of triage calls. The average caller is a near-elderly patient (average age = 63.1) with high rates of healthcare utilization in the previous year: 97.4% for primary care, 38.2% for ED visits, and 10.2% for hospital admissions. Nurses recommend primary care (PC) for 62.7% of calls, urgent care (UC) for 4.4%, and emergency department (ED) care for 27.7%. Utilization rates within three days of the call are low: 30.7% of patients visit PC, 6.3% visit UC, and 19.2% visit the ED. Only 49.9% receive any form of care (PC, UC, or ED) within three days.

4. Empirical methods

4.1. Overview

Our primary goal is to analyze how nurse recommendations influence patient healthcare choices after the triage call. For simplicity, suppose that the patient has four alternative options: (i) self-care (or no care) at home ($Home$), (ii) primary care (PC), (iii) urgent care (UC), and (iv) emergency department (ED). Likewise, suppose that the nurse selects one of the four follow-up care recommendations: $Home$, PC, UC, or ED. We denote the joint distribution of the patient's healthcare utilization as follows:

$$F(Home_{id}, PC_{id}, UC_{id}, ED_{id} \mid R_i) \tag{1}$$

where ED_{id} indicates whether patient i visits an ED within d days of triage. UC_{id}, PC_{id}, and $Home_{id}$ are indicators of urgent care, primary care, and self-care at home, respectively. $R_i \in \{Home, PC, UC, ED\}$ indicates the health care option recommended by the nurse.

[7] We impose this restriction since our identification strategy relies on cross-nurse variation within cells defined by algorithm disposition as described in Section 4.2. This restriction ensures that nurses' ED tendencies are estimated with a sufficiently large number of observations. Anecdotally, some nurse managers stated that some nurses would only work nurse triage for short periods, or that nurse managers themselves would occasionally step in to field calls when needed.

Table 2
Effect of nurse ED vs. UC recommendation.

(a) IV

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends ED	2.662	−24.025	25.342	2.506	4.137	1.603	0.272	110.745	104.120	214.865
	(0.859)	(0.637)	(0.816)	(0.965)	(0.984)	(0.355)	(0.294)	(35.548)	(38.501)	(62.305)
Outcome mean (UC)	19.588	38.646	11.392	48.278	58.877	1.214	2.292	227.906	93.061	320.967
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends ED	1.267	−24.587	24.203	2.196	0.930	2.198	−1.413	160.101	243.998	404.099
	(0.966)	(0.669)	(0.863)	(0.979)	(0.839)	(0.450)	(0.580)	(82.975)	(132.806)	(176.258)
Outcome mean (UC)	52.305	42.171	18.192	55.688	77.640	2.788	9.790	1217.306	546.884	1764.190
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

(b) OLS

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends ED	1.750	−31.736	35.534	5.024	4.411	3.950	−0.051	183.860	231.542	415.402
	(0.257)	(0.252)	(0.255)	(0.316)	(0.304)	(0.103)	(0.091)	(10.860)	(11.009)	(18.894)
Outcome mean (UC)	19.588	38.646	11.392	48.278	58.877	1.214	2.292	227.906	93.061	320.967
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends ED	4.983	−32.336	33.494	3.974	3.454	4.591	−0.186	420.264	797.949	1218.212
	(0.299)	(0.255)	(0.272)	(0.312)	(0.249)	(0.131)	(0.181)	(24.166)	(39.146)	(52.353)
Outcome mean (UC)	52.305	42.171	18.192	55.688	77.640	2.788	9.790	1217.306	546.884	1764.190
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

Notes: This table reports TSLS estimates of the effect of ED recommendation. All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. In the third row of each panel, the sample average for each outcome conditional on UC recommendation is reported as a benchmark ($E\left[Y_i \mid R_i = UC\right]$). In the fourth row, the overall (unconditional) sample average is also reported ($E\left[Y_i\right]$). In addition to the baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), all regressions control for the hold-out variables (age, sex, marital status, race/ethnicity, period of service, prior healthcare utilization, and prior diagnoses) for precision. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. Standard errors are clustered at the call center-by-call time level.

We would like to contrast this conditional distribution (1) evaluated at different recommendations $R_i \in \{Home, PC, UC, ED\}$. For example, consider a case where the nurse changes the triage recommendation from UC to ED. If recommendations were randomly determined, we would know the effect of this change on the patient choice probability by contrasting two distributions:

$$\begin{aligned} &F(Home_{id}, PC_{id}, UC_{id}, ED_{id} \mid R_i = ED) \\ &\quad \text{vs.} \\ &F(Home_{id}, PC_{id}, UC_{id}, ED_{id} \mid R_i = UC). \end{aligned} \tag{2}$$

In general, it is difficult to identify the effect of nurse recommendation with observational data, as the nurse does not select a follow-up care recommendation at random. The nurse is more likely to recommend an acute form of care when patients present more acute and severe health conditions. Patients with acute symptoms may use an acute form of care, regardless of the nurse's recommendations. Hence, the contrast (2) may only tell us differences in health conditions between patients when it is calculated with observational data.

We address this identification problem by exploiting the quasi-random assignment of triage nurses to calls and variations in triage recommendations across those nurses. We exploit that different nurses have different preferences on the appropriate level of follow-up care even for patients with the same health conditions, analogous to the findings in the literature on physicians' practice variations.[8]

The following subsections describe our research design. Although we analyze margin-specific causal effects for each pair of the two adjacent triage recommendations (ED-UC, UC-PC, and PC-Home), the following sections describe our empirical strategy for the ED-UC margin for compactness. Appendix Sections A1, A2, and A3 detail our strategy. In what follows, we use "call" and "patient" interchangeably for notational simplicity unless otherwise noted, although the unit of observation i is a call.

4.2. Measuring cross-nurse variation in triage recommendations

For each call i received by call center c at time t, we observe the assigned nurse J_i and the recommended care location $R_i \in \{Home, PC, UC, ED\}$. To derive an empirical model, we convert the recommendation status R_i into a set of four indicators

[8] Some examples include Coussens and Ly (2025), Abualenain et al. (2013), and Molitor (2018).

Table 3
Effect of nurse UC vs. PC recommendation.

(a) IV

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends UC	−11.484	20.693	−1.225	18.634	4.812	−0.091	−0.367	−69.782	−23.403	−93.185
	(0.544)	(0.464)	(0.452)	(0.589)	(0.622)	(0.178)	(0.183)	(19.522)	(22.351)	(35.774)
Outcome mean (PC)	35.228	5.306	8.677	13.746	45.423	0.973	2.115	225.842	52.582	278.425
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends UC	−7.884	21.009	−1.348	17.447	2.140	−0.382	−0.040	−137.188	−109.740	−246.929
	(0.604)	(0.482)	(0.516)	(0.608)	(0.538)	(0.243)	(0.361)	(48.074)	(72.106)	(99.177)
Outcome mean (PC)	65.879	6.956	15.602	21.731	73.368	2.656	10.225	1261.896	530.963	1792.860
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

(b) OLS

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends UC	−14.007	29.177	2.378	30.276	11.972	0.266	0.195	−23.032	14.673	−8.359
	(0.219)	(0.241)	(0.159)	(0.262)	(0.250)	(0.056)	(0.073)	(6.010)	(6.795)	(11.099)
Outcome mean (PC)	35.228	5.306	8.677	13.746	45.423	0.973	2.115	225.842	52.582	278.425
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends UC	−11.728	29.527	2.430	28.893	4.127	0.182	0.158	−87.873	21.901	−65.972
	(0.247)	(0.241)	(0.189)	(0.259)	(0.208)	(0.082)	(0.143)	(16.386)	(24.554)	(33.756)
Outcome mean (PC)	65.879	6.956	15.602	21.731	73.368	2.656	10.225	1261.896	530.963	1792.860
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

Notes: This table reports TSLS estimates of the effect of UC recommendation. All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. In the third row of each panel, the sample average for each outcome conditional on PC recommendation is reported as a benchmark ($E\left[Y_i \mid R_i = PC\right]$). In the fourth row, the overall (unconditional) sample average is also reported ($E\left[Y_i\right]$). In addition to the baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), all regressions control for the hold-out variables (age, sex, marital status, race/ethnicity, period of service, prior healthcare utilization, and prior diagnoses) for precision. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. Standard errors are clustered at the call center-by-call time level.

$R_i^r = 1\{R_i = r\}$, $r \in \{Home, PC, UC, ED\}$. To capture cross-nurse variation in triage practice, we denote each nurse j's propensity to recommend option r by $Z_j^r = E\left[R = r \mid J = j\right]$, $r \in \{Home, PC, UC, ED\}$. The probability vector $\left(Z_j^{ED}, Z_j^{UC}, Z_j^{PC}, Z_j^{Home}\right)$ characterizes nurse j's triage style, and the four elements sum to one.

Empirically, we construct a leave-one-out instrument by averaging ED recommendation indicators of other patients triaged by the same nurse, following the examiner design literature. Specifically, for call i that is assigned to nurse j, we first obtain residual of ED recommendation status, denoted as R_i^{ED*}, before calculating the leave-one-out average, using the following linear regression:

$$R_i^{ED*} = R_i^{ED} - X_i^{0\prime}\Gamma_x - H_i^{\prime}\Gamma_h = Z_{ij}^{ED} + \epsilon_i \tag{3}$$

where X_i^0 is a vector of our baseline controls, including interactions of call center-by-call time (month-year, day-of-week, and AM/PM indicators), interactions of call center-by-algorithm disposition, and county of patient's residence (5-digit FIPS indicators). In addition to X_i^0, we also partial out a vector of additional controls H_i ("hold-out controls"), including age, sex, marital status, race and ethnicity, period of service, prior healthcare utilization, and prior diagnoses. Importantly, the hold-out set H_i is not essential for balance, but is included for statistical precision.[9] The residuals R_i^{ED*} include nurse j's ED tendency Z_{ij}^{ED} and idiosyncratic call-level error term ϵ_i.

Then we construct the leave-one-out measure for call i by averaging the residual ED recommendation of all other patients but patient $k(i)$ assigned to nurse j in call center c with algorithm disposition a:

$$Z_i^{ED} = \frac{1}{K_{jcya} - 1} \sum_{i'} \frac{1\left\{k(i') \neq k(i), j(i') = j, c(i') = c, y(i') = y, a(i') = a\right\} R_{i'}^{ED*}}{n_{k(i')jcya}} \tag{4}$$

where K_{jcya} is the number of patients assigned to nurse j in call center c in year y with algorithm disposition a and n_{kjcya} is the total number of calls from patient k received by nurse j in call center c in year y with algorithm disposition a. We call this Z_i^{ED} as the leave-one-patient-out measure and use it as an instrument for whether patient i is recommended ED or UC.[10] We similarly construct leave-out measures of the nurse's propensity to recommend UC, PC, and Home (Z_i^{UC}, Z_i^{PC}, Z_i^{Home}).[11]

[9] We check balance without controlling for the hold-out set H_i. We confirm that patients are balanced between nurses within the baseline controls X_i^0.

4.3. Main empirical specification

We exploit cross-nurse variation in recommending ED over UC (Z_i^{ED}) as an exogenous shock to whether the patient is recommended ED or UC (R_i^{ED}), holding constant cross-nurse variation in non-focal recommendation tendencies (Z_i^{PC}, Z_i^{Home}). Intuitively, this method focuses on nurses who have the same tendencies to recommend PC and Home, and among those nurses, we compare patients assigned to high-ED (low-UC) nurses to patients assigned to low-ED (high-UC) nurses.

Our first-stage equation explains patient i's recommendation status (ED or UC) by the assigned nurse's triage style Z_i's, patient observable characteristics $\left(X_i^0, H_i\right)$, and idiosyncratic shock v_i:

$$R_i^{ED} = \alpha_0 + \alpha_1 Z_i^{ED} + \alpha_2 Z_i^{PC} + \alpha_3 Z_i^{Home} + X_i^{0\prime} \delta_x + H_i' \delta_h + v_i \tag{5}$$

where Z_i^{ED} is the focal instrument with Z_i^{UC} being left out of the equation as the reference category. The non-focal instruments Z_i^{PC} and Z_i^{Home} are included as controls to make our comparison only based on cross-nurse variation in ED and UC tendencies.

In our second stage equation, we examine the effect of being recommended ED over UC on several patient outcomes Y_i within 3 to 28 days of triage, including health care utilization (primary care, urgent care, ED, hospital admission), costs (outpatient, inpatient, and total costs), and repeat calls. Using the predicted ED recommendation status $\hat{R}_i^{ED}$ as an instrument, we estimate the effect of being recommended ED over UC as follows:

$$Y_i = \beta_0 + \beta_1 \hat{R}_i^{ED} + \beta_2 Z_i^{PC} + \beta_3 Z_i^{Home} + X_i^{0\prime} \pi_x + H_i' \pi_h + u_i \tag{6}$$

where β_1 is our parameter of interest that captures the effect of ED recommendation, relative to UC recommendation. This specification controls for the non-focal propensities Z_i^{PC} and Z_i^{Home} to estimate the margin-specific LATE among the patients on the ED vs. UC recommendation margin.[12]

4.4. Assessing the identifying assumption

In this section, we briefly discuss and assess the assumption for our IV strategy to identify the LATE of being recommended ED over UC. For curious readers, Appendix Sections A1 and A3 more formally describe the IV condition that allows us to identify the margin-specific LATE, following Humphries et al. (2024). An additional monotonicity test is also presented in Appendix Section A1. Appendix Section A2 further shows a behavioral model of nurse decision-making that aligns with our IV assumption.

4.4.1. Relevance

We begin our empirical analysis by examining each of the IV assumptions before presenting our TSLS estimation results. First, we test for instrument relevance. This assumption requires that the leave-out measure of ED recommendations Z_i^{ED} meaningfully predicts the actual ED recommendation R_i^{ED} ($\alpha_1 > 0$), conditional on the leave-out PC and Home measures, Z_i^{PC} and Z_i^{Home}, and the baseline controls X_i^0.

Fig. 1(a) presents a binned scatter plot of residualized ED recommendation R_i^{ED} on the y-axis against residualized leave-out measure Z_i^{ED} on the x-axis. The figure indicates that shifting the call assignment from the 5 percentile (leftmost) bin to the 95 percentile (rightmost) increases the probability of receiving an ED recommendation from 15.8% to 39.3%, holding the leave-out PC measure constant. The solid line represents the estimated first stage coefficient $\hat{\alpha}_1 = 0.8595$ (SE = 0.0067), confirming the first-stage relationship between R_i^{ED} and Z_i^{ED} is statistically significant.

Similarly, Fig. 1(b) presents the first-stage relationship between R_i^{UC} and Z_i^{UC}, conditional on the leave-out ED and Home propensities Z_i^{ED} and Z_i^{Home} ($\hat{\alpha}_1$ = 0.9196, SE = 0.0050). Likewise, Fig. 1(c) shows the first-stage relationship between R_i^{PC} and Z_i^{PC}, conditional on the leave-out ED and UC propensities Z_i^{ED} and Z_i^{UC} ($\hat{\alpha}_1 = 0.9376$, SE = 0.0057).

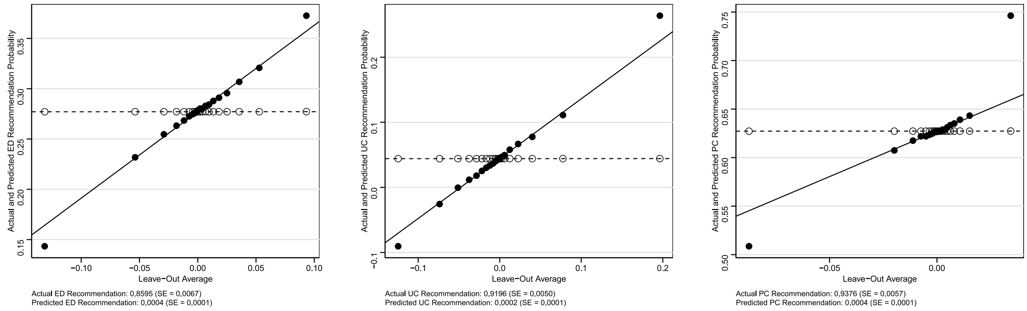
4.4.2. Conditional independence and exclusion

We next consider the conditional independence assumption. We consider this assumption to be reasonable. The assignment of telephone triage nurses to patients is as good as random since incoming calls are allocated to the next available nurse within call centers, and neither patients nor nurses are able to select whom they speak with.

[10] We construct our leave-out ED propensity within cells defined by call center, call year, and algorithm disposition (follow-up location and timing) to accommodate that the nurse's ED tendency relative to other nurses likely vary across different cells. This method allows us to mitigate concerns over monotonicity violation (Mueller-Smith, 2015; Sigstad, 2023). See Appendix Section A5 for further details.

[11] As robustness checks, we implement our IV design by constructing the leave-one-out measures in several different ways and confirm that the results are unaffected. Appendix Section A6 details our alternative measures.

[12] An alternative specification uses a second-stage equation with the three treatments, R_i^{ED}, R_i^{PC}, and R_i^{Home}, instrumented with the three triage style measures, Z_i^{ED}, Z_i^{PC}, and Z_i^{Home}. Under the IV assumptions, this specification produces the same TSLS estimand as Eqs. (5) and (6) (Humphries et al., 2024). Our OLS results come from this specification, with the three treatments not instrumented with the nurse triage measures.



(a) ED - UC. (b) UC - PC. (c) PC - Home.

Fig. 1. First stage and balance.
Notes: The binned scatter plot in Panel (a) represents the first-stage regression in Eq. (5) and the corresponding balance regression that replaces actual ED recommendation with predicted ED recommendation. The bin averages of the actual ED recommendation R_i^{ED} are shown in solid circles, whereas the bin averages of the predicted ED recommendation $\hat{R}_i^{ED}$ are shown in hollow circles. Both R_i^{ED} and $\hat{R}_i^{ED}$ are residualized for the baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), and are centered at the average ED recommendation rate (27.7%) on the y-axis. Those baseline controls are described in Appendix Table B3. Panels (b) and (c) represent the corresponding analysis on UC-PC and PC-Home margins.

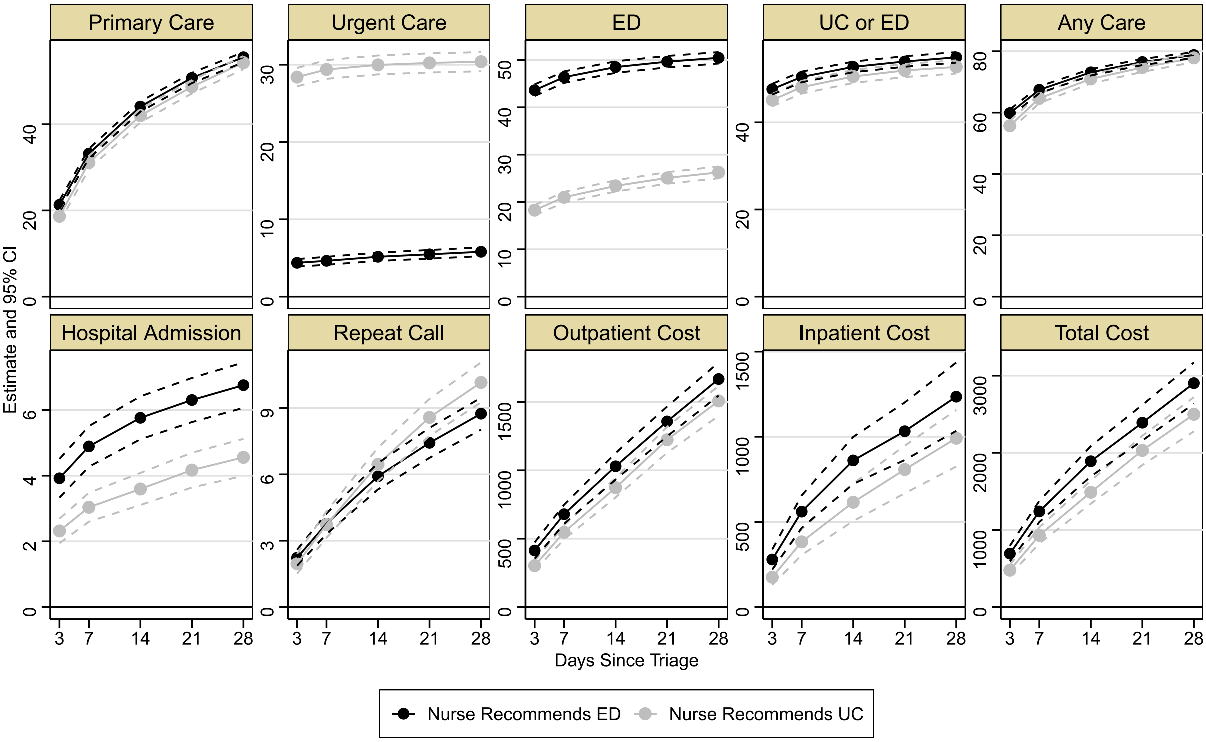


Fig. 2. Average potential outcomes among compliers (ED - UC).
Notes: This figure presents the averages of potential outcomes among compliers. The black circles are the counterfactuals under ED recommendation ($E\left[Y(ED) \mid \text{Compliers}\right]$), while the gray circles are under UC recommendation ($E\left[Y(UC) \mid \text{Compliers}\right]$). All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. Following Abadie (2003), we estimate $E\left[Y(ED) \mid \text{Compliers}\right]$ by regressing an interaction between each outcome and ED recommendation indicator ($Y \cdot R^{ED}$) on ED recommendation indicator (R^{ED}) with the right-hand-side R^{ED} instrumented by the leave-out measure. $E\left[Y(UC) \mid \text{Compliers}\right]$ is similarly estimated by replacing R^{ED} with $1 - R^{ED}$. In addition to the baseline controls X_i^0, all regressions control for the hold-out variables. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. The dashed lines show the 95% confidence intervals based on standard errors clustered at the call center-by-call time level.

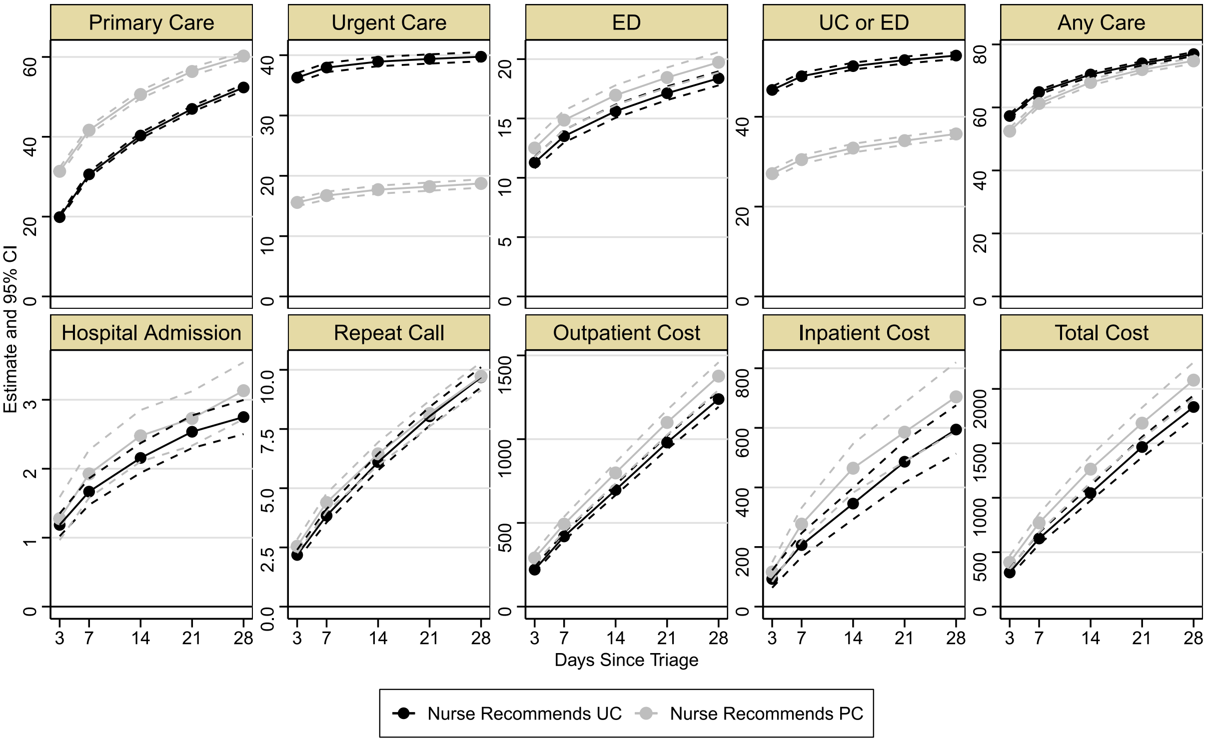


Fig. 3. Average potential outcomes among compliers (UC - PC).
Notes: This figure presents the averages of potential outcomes among compliers. The black circles are the counterfactuals under UC recommendation ($E\left[Y(UC) \mid \text{Compliers}\right]$), while the gray circles are under PC recommendation ($E\left[Y(PC) \mid \text{Compliers}\right]$). All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. Following Abadie (2003), we estimate $E\left[Y(UC) \mid \text{Compliers}\right]$ by regressing an interaction between each outcome and UC recommendation indicator ($Y \cdot R^{UC}$) on UC recommendation indicator (R^{UC}) with the right-hand-side R^{UC} instrumented by the leave-out measure. $E\left[Y(PC) \mid \text{Compliers}\right]$ is similarly estimated by replacing R^{UC} with $1 - R^{UC}$. In addition to the baseline controls X_i^0, all regressions control for the hold-out variables. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. The dashed lines show the 95% confidence intervals based on standard errors clustered at the call center-by-call time level.

We empirically examine conditional independence by testing whether the leave-out ED propensity Z_i^{ED} is correlated with patient hold-out characteristics H_i. Following standard balance checks in the examiner design literature, we first estimate each call's predicted ED recommendation probability, $\widehat{R}_i^{ED}$, by regressing ED recommendation indicator, R_i^{ED}, on baseline controls, X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), and hold-out controls, H_i (age, sex, marital status, race and ethnicity, period of service, prior healthcare utilization, and prior diagnose). Next, we test whether $\widehat{R}_i^{ED}$ is correlated with Z_i^{ED}, conditional on X_i^0, Z_i^{PC}, and Z_i^{Home}.

Fig. 1 shows that the leave-out measure, Z_i^{ED}, is not meaningfully related to patient characteristics, as indicated by the flat dashed line. We dig into this deeper in Appendix Figure B3, which presents a multivariate regression of Z_i^{ED} on hold-out controls, H_i, conditional on X_i^0. While the balance test fails on certain variables – notably age – the level differences are small at about 0.1 percent. More importantly, unlike in the right panel that shows the estimated coefficients of actual ED recommendation on hold-out controls, the F-test fails to reject the null at the 10% significance level, suggesting that hold-out variables H_i do not jointly predict nurses' ED recommendation tendency. Likewise, Appendix Figure B4 show that nurses' UC propensity is not meaningfully related to hold-out characteristics. For nurses' PC propensity in Appendix Figure B5, although the F-test is marginally significant we show differences in individual characteristics are similarly small.

Finally, the exclusion restriction requires that there not be multiple channels by which the instrument affects outcomes. While it would be advantageous to identify the effect of ED care over another type of care directly, multiple treatment options renders this analytically difficult. First-stage results, available on request, of a parallelly constructed leave-one-out instrument of nurse-specific patient ED probability shows that being assigned a high-ED nurse moves not just a patient's ED probability but their PC probability as well. Lacking additional instruments, we consider receiving a given recommendation as the first stage as described above to satisfy this requirement.

4.4.3. Monotonicity

We assess the monotonicity assumptions to identify interpretable average treatment effects in the presence of heterogeneous treatment effects. In our context, monotonicity requires that, for any pair of nurses j, j' with the same PC and Home propensities

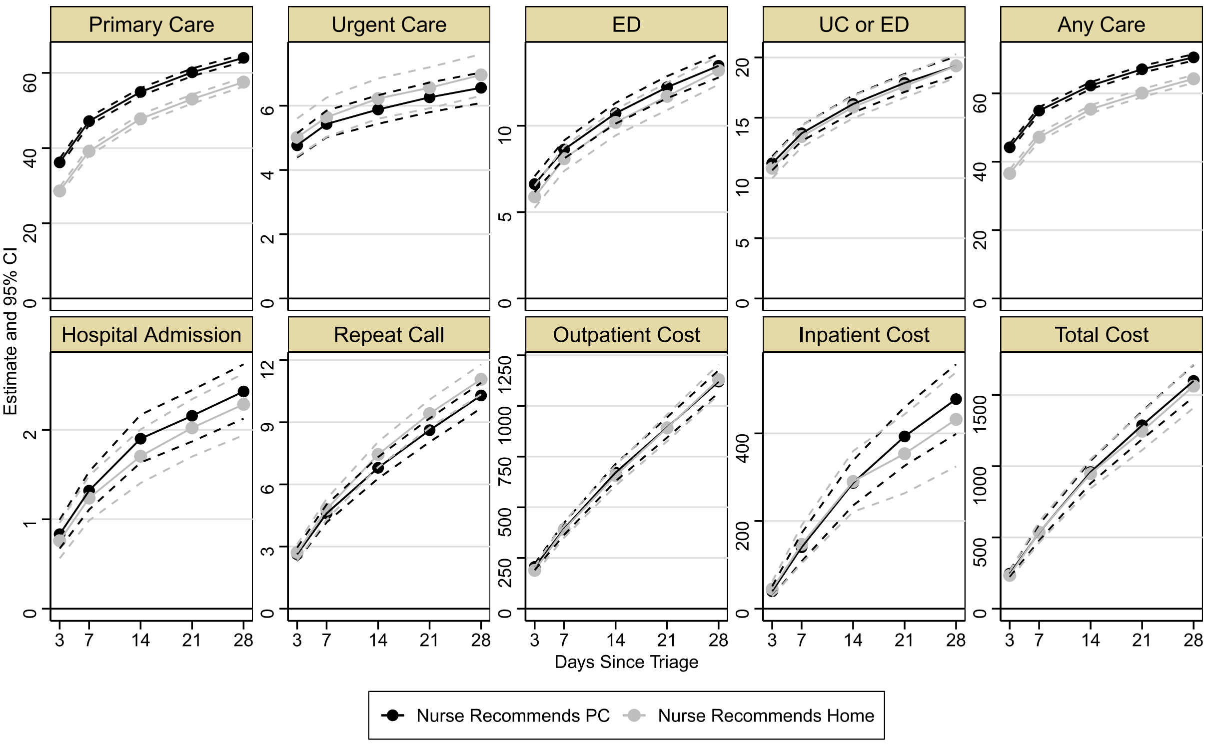


Fig. 4. Average potential outcomes among compliers (PC - Home).
Notes: This figure presents the averages of potential outcomes among compliers. The black circles are the counterfactuals under UC recommendation ($E\left[Y(PC) \mid \text{Compliers}\right]$), while the gray circles are under PC recommendation ($E\left[Y(Home) \mid \text{Compliers}\right]$). All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. Following Abadie (2003), we estimate $E\left[Y(PC) \mid \text{Compliers}\right]$ by regressing an interaction between each outcome and PC recommendation indicator ($Y \cdot R^{PC}$) on PC recommendation indicator (R^{PC}) with the right-hand-side R^{PC} instrumented by the leave-out measure. $E\left[Y(Home) \mid \text{Compliers}\right]$ is similarly estimated by replacing R^{PC} with $1 - R^{PC}$. In addition to the baseline controls X_i^0, all regressions control for the hold-out variables. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. The dashed lines show the 95% confidence intervals based on standard errors clustered at the call center-by-call time level.

(z^{PC}, z^{Home}), if nurse j' has a higher ED propensity than nurse j ($z'^{ED} > z^{ED}$), then all calls that would be recommended ED by nurse j must be also recommended ED by nurse j' ($R_i^{ED}(z'^{ED}, z^{PC}, z^{Home}) \geq R_i^{ED}(z^{ED}, z^{PC}, z^{Home})$). This assumption may fail for several potential reasons. For instance, nurse j' may have a lower ED propensity than nurse j for younger patients ($R_i^{ED}(z'^{ED}, z^{PC}, z^{Home}) = 0 < 1 = R_i^{ED}(z^{ED}, z^{PC}, z^{Home}) \mid Age < 65$), even when nurse j' has a higher ED propensity than nurse j overall.

To empirically assess monotonicity, we implement a conventional test for (average) monotonicity by examining whether the first-stage coefficients are positive across subsamples defined by observable characteristics (Frandsen et al., 2023). Specifically, we split the sample based on several characteristics of the patients and estimate the coefficients of the first stage separately. Appendix Tables B5–B7 show that the leave-out ED (/UC/PC) propensity is positively associated with ED (/UC/PC) recommendation for all subsamples. In Section 5.4, we further show that the potential response estimates for binary outcomes are bounded between 0 and 1, suggesting stronger support for monotonicity.[13]

5. Results

This section presents TSLS estimates of margin-specific causal effects of triage recommendations on patient utilization, health, and costs. Naive OLS are presented alongside for comparison. We present results in separate tables and figures by recommendation margins (ED-UC, UC-PC, PC-Home). All outcomes are cumulative from day 1 (triage call day) through day x ($x = 3, \ldots, 28$).

[13] Chan et al. (2022) shows that this potential outcome test is stronger than the conventional test for (average) monotonicity, following Kitagawa (2015).

5.1. Utilization

Table 2 and Appendix Figure B6 present the TSLS estimates from Eqs. (6) and (5), capturing margin-specific causal effects of being recommended ED over UC among compliers—those for whom nurses vary in their recommendation between ED and UC. The TSLS estimates show that the ED recommendation significantly shifts patients' utilization from UC to ED. Patients recommended ED care are 24.0 percentage points less likely to use UC and 25.3 percentage points more likely to use ED within three days of the call, compared to patients recommended UC. This makes them more than three times more likely to go to the ED as the average caller recommended UC care, and more than twice more likely as the overall sample mean. With the large increase in ED utilization, we observe a rise in emergency admissions through the ED. Patients recommended ED care are 1.6 percentage points more likely to experience an emergency admission than those recommended UC, a 133 percent increase. These estimates are generally more muted than corresponding naive OLS estimates, which show a higher probability of an ED visit and hospital admission.

The ED recommendation modestly shifts patient utilization beyond the corresponding care margin. Patients recommended ED are 2.7 percentage points (13.7 percent) more likely to go to primary care in three days, compared to patients recommended UC. We also find that the ED recommendation increases both the probability of receiving any acute care (UC or ED) and any form of professional care (PC, UC, or ED) modestly, by 2.5 percentage points (5.2 percent) and 4.1 percentage points (6.9 percent), respectively.

In terms of the timing of the effect, overall, the TSLS estimates suggest that the ED recommendation influences patient utilization only in the very short run on day 1 through day 3. The TSLS estimates for UC and ED utilization on day 28 (−24.6 p.p. and +24.2 p.p., respectively) do not differ much from the estimates on day 3 (−24.0 p.p. and +25.3 p.p., respectively). This implies that the changes in patient UC and ED utilization mostly occur on day 1 through day 3, and then there is no additional recommendation effect thereafter.[14] By contrast, the effect on any care utilization shrinks and becomes insignificant on day 28.

Table 3 and Appendix Figure B7 show estimates for patients on the middle acute margin between urgent care and primary care recommendations. Unlike the ED-UC margin, the estimated effects on UC and PC utilization are not of a similar magnitude. Patients recommended UC care are 20.7 percentage points more likely to go to UC and 11.5 percentage points less likely to use primary care within three days of the triage call. These patients are also 1.2 percentage points less likely to use the ED in this time period, which translates 13.8 percent compared to the average ED utilization rate of patients recommended PC. The UC recommendation modestly increases the probability of receiving any professional care by 4.8 percentage points (10.6 percent). We find no difference in the probability of a hospital admission.

Table 4 and Appendix Figure B8 show estimates for patients on the least acute margin between PC and Home recommendations. The TSLS estimates suggest that the PC recommendation changes patient utilization almost exclusively on the corresponding care margin of PC and Home. Patients recommended PC are 7.6 percentage points (27.8 percent) more likely to use primary care within three days, compared to patients recommended Home. The effects on any care are 7.6 percentage points on day 3 and 6.3 percentage points on day 28. The effects on urgent care, ED utilization, and hospital admission are small and statistically insignificant.

We next examine if patients call the triage line again in a short period of time as a proxy for not having the health issue resolved. Table 2 and Appendix Figure B6 show that an ED recommendation reduces repeat calls within 28 days of the initial call by 1.4 p.p. relative to a sample mean of 10.3 percent. In contrast, we do not see a statistically significant reduction than those on the moderately acute UC-PC margin (−0.4 p.p. (0.3), Table 3 and Appendix Figure B7) or the PC and home margin (−0.78 p.p. (0.48), Table 4 and Appendix Figure B8). This may indicate that the ED visit is more likely to resolve the issue for the patient. We further find that although the ED recommendation substantially shifts patient utilization toward more intensive care, we do not find evidence that patients recommended ED experience any survival gain, compared to patients recommended UC. We also find no effect on other acuity margins. However, short-term mortality is a relatively rare event for this population (0.4 percent within 28 days) and these estimates are imprecise.[15]

5.2. Costs

In Table 2, we find that, on average, patients recommended ED care incur \$110.7 higher outpatient costs than those recommended UC within three days. Although this net increase in outpatient costs can be driven by both the large shift in utilization from UC to ED and the modest shift from self-care to any professional outpatient care (PC, UC, or ED), Appendix Table B11a suggests that a large part of the increased costs is likely explained by the higher cost of an ED visit without admission compared to a UC visit. Additionally, Table 2 shows that patients recommended ED incurs \$104.1 in additional inpatient costs within three days, compared to those recommended UC, consistent with a 1.6 percentage point higher hospital admission rate. Altogether, patients recommended ED care over UC care incur 22.9 percent (\$404) higher cumulative costs after 28 days.

Interestingly, for patients on the UC-PC recommendation margin, Table 3 and Appendix Figure B7 show that patients recommended UC (more acute option) incur *lower* costs, compared to patients recommended PC (less acute option). Outpatient costs for patients recommended UC are \$69.8 (30.1 percent) lower after three days. In Appendix Table B11b, we find that the UC recommendation decreases primary care costs (−\$81.7) *more* than it increases urgent care costs (+\$52.6), although the UC recommendation decreases primary care utilization (−11.5 p.p.) *less* than it increases urgent care utilization (+20.7 p.p.).[16]

[14] We confirm this by estimating the average potential outcomes under each counterfactual recommendation in Section 5.4.

[15] See Appendix Tables B12, B13, and B14.

[16] This finding can be explained by differences in the type and number of services provided. An urgent care visit tends to be more focused on addressing the chief complaint, while a primary care visit potentially addresses multiple concerns beyond the chief complaint. This difference can make an urgent care visit less costly than a primary care visit.

Table 4
Effect of nurse PC vs. Home recommendation.

(a) IV

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends PC	7.598	−0.249	0.738	0.414	7.627	0.072	−0.088	16.921	−5.174	11.747
	(0.718)	(0.348)	(0.398)	(0.508)	(0.769)	(0.131)	(0.255)	(11.940)	(10.491)	(19.226)
Outcome mean (Home)	27.277	3.447	4.937	8.276	33.401	0.536	2.917	164.209	26.862	191.072
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends PC	6.443	−0.398	0.265	−0.007	6.259	0.144	−0.780	−9.933	46.234	36.301
	(0.778)	(0.398)	(0.531)	(0.621)	(0.744)	(0.232)	(0.479)	(46.162)	(67.379)	(93.704)
Outcome mean (Home)	56.330	5.163	11.749	16.352	62.204	2.130	11.728	1100.355	463.917	1564.272
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

(b) OLS

	Primary care	Urgent care	ED	UC or ED	Any care	Hospital admission	Repeat call	Outpatient cost	Inpatient cost	Total cost
Panel A: 3-day outcomes										
Nurse recommends PC	11.226	0.400	1.089	1.422	12.018	0.190	−0.577	39.454	3.553	43.007
	(0.386)	(0.178)	(0.194)	(0.254)	(0.411)	(0.068)	(0.134)	(6.309)	(5.399)	(10.168)
Outcome mean (Home)	27.277	3.447	4.937	8.276	33.401	0.536	2.917	164.209	26.862	191.072
Outcome mean (All)	30.672	6.311	19.187	25.159	49.857	2.671	2.251	290.885	157.584	448.469
Panel B: 28-day outcomes										
Nurse recommends PC	11.852	0.315	0.714	1.050	11.736	0.151	−1.349	82.935	27.984	110.919
	(0.419)	(0.207)	(0.269)	(0.319)	(0.403)	(0.121)	(0.258)	(26.011)	(34.633)	(49.636)
Outcome mean (Home)	56.330	5.163	11.749	16.352	62.204	2.130	11.728	1100.355	463.917	1564.272
Outcome mean (All)	63.081	7.977	25.912	32.782	74.960	4.668	10.299	1406.952	855.097	2262.050

Notes: This table reports TSLS estimates of the effect of PC recommendation. All binary outcomes (primary care - mortality) are multiplied by 100 so that estimates can be interpreted as percentage points. Costs are measured in US dollars. In the third row of each panel, the sample average for each outcome conditional on Home recommendation is reported as a benchmark ($E\left[Y_i \mid R_i = Home\right]$). In the fourth row, the overall (unconditional) sample average is also reported ($E\left[Y_i\right]$). In addition to the baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), all regressions control for the hold-out variables (age, sex, marital status, race/ethnicity, period of service, prior healthcare utilization, and prior diagnoses) for precision. The baseline and hold-out controls, X_i^0 and H_i, are described in Appendix Tables B3 and B4, respectively. Standard errors are clustered at the call center-by-call time level.

Table 5
Complier characteristics.

Variable	Overall mean	Compliers (ED-UC)		Compliers (UC-PC)		Compliers (PC-Home)	
		Mean	Ratio	Mean	Ratio	Mean	Ratio
Comorbidity count	2.901	3.003	1.035	2.742	0.945	2.628	0.906
		(0.025)	[1.018–1.052]	(0.013)	[0.936–0.954]	(0.018)	[0.894–0.918]
Age	63.054	63.646	1.009	60.772	0.964	60.477	0.959
		(0.183)	[1.004–1.015]	(0.118)	[0.960–0.967]	(0.162)	[0.954–0.964]
Male	0.863	0.868	1.005	0.828	0.959	0.850	0.985
		(0.004)	[0.996–1.014]	(0.003)	[0.953–0.966]	(0.004)	[0.977–0.993]
Married	0.513	0.514	1.003	0.505	0.986	0.501	0.977
		(0.006)	[0.979–1.026]	(0.004)	[0.971–1.000]	(0.005)	[0.958–0.996]
White	0.691	0.679	0.982	0.607	0.878	0.655	0.948
		(0.006)	[0.966–0.999]	(0.004)	[0.866–0.890]	(0.005)	[0.933–0.962]
Black	0.182	0.181	0.993	0.239	1.312	0.190	1.043
		(0.005)	[0.941–1.044]	(0.004)	[1.272–1.353]	(0.004)	[0.999–1.087]
Hispanic	0.061	0.073	1.196	0.086	1.424	0.087	1.441
		(0.003)	[1.092–1.301]	(0.003)	[1.336–1.512]	(0.003)	[1.335–1.547]

Notes: This table presents the average characteristics among compliers, along with the overall sample averages. Following Abadie (2003), we estimate $E\left[H_i \mid \text{Compliers}\right]$ by regressing an interaction between each patient characteristic H and ED recommendation indicator ($H \cdot R^{ED}$) on ED recommendation indicator (R^{ED}) with the right-hand-side R^{ED} instrumented by the leave-out ED recommendation propensity on ED-UC margin. The average characteristics among compliers on UC-PC and PC-Home margins are estimated similarly. All regressions control for the baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county) described in Appendix Table B3. Standard errors are clustered at the call center-by-call time level.

For patients on the PC-Home recommendation margin, we do not find significant differences in costs between patients recommended PC and those recommended Home in Table 4 and Appendix Figure B8. 28 days after the call, patients recommended PC incur only \$36 higher total costs than patients recommended Home, and the effect is statistically insignificant. The null effects on costs are consistent with our finding that the PC recommendation only has a small impact on the probability of having a primary care visit.

5.3. Robustness checks

We examine the sensitivity of the TSLS estimates to the choice of control variables. Although our balance check shows that our leave-out measure is uncorrelated with patients' predetermined observable characteristics after controlling for baseline controls X_i^0, it may still be correlated with patients' potential outcomes through unobservable factors. To assess the severity of this threat to conditional independence, we test the sensitivity of our TSLS estimates to (i) the exclusion of hold-out controls from the conditioning set[17] and (ii) the inclusion of several symptomatic measures recorded at the time of the call. Appendix Table B15 reports estimates from the specification that controls only for the baseline control X_i^0. Appendix Table B16 presents TSLS estimates that additionally control for pain scale (0–10) and duration of symptoms (in 10 bins). Lastly, Appendix Table B17 shows TSLS results from the specification that further controls for chief complaint fixed effects. The TSLS estimates remain generally stable across different sets of controls.

Appendix Section A6 further examines the robustness of the TSLS estimates to alternative leave-out measures. Again, we confirm that our findings are not sensitive to how we construct these measures.[18]

5.4. Evaluation of mean potential outcomes

While the TSLS estimates recover margin-specific LATEs among compliers, these causal effect parameters do not provide insight into how the levels of potential outcomes evolve over time under each of the counterfactual recommendations. In Figs. 2–4, we further illustrate evolutions of the average potential outcomes for compliers when recommended the more intensive option for care (black circles, ED when on ED-UC margin, UC when on UC-PC margin, and PC when on PC-Home margin) and when recommended the less intensive option (gray circles, UC when on ED-UC margin, PC when on UC-PC margin, and Home when on PC-Home margin).[19] Note that, in each panel, the gaps in estimated potential outcomes between the two counterfactual recommendations are numerically equivalent to our TSLS estimates.

As discussed in Section 5.1, for patients on ED-UC recommendation margin, our TSLS estimates suggest that the ED recommendation shifts patient utilization from UC to ED immediately after the call (day 1 through day 3) and then the recommendation has no additional effect thereafter. Fig. 2 confirms this finding: the average potential ED (UC) utilization rates under the two counterfactual recommendations run in parallel from day 3 to day 28, with an almost constant difference (25 percentage points). By contrast, the differences in potential utilization of any care (PC, UC, or ED) between the two counterfactuals shrink over time, suggesting that the ED recommendation may modestly shift the timing of having any care. We find similar patterns in potential utilization for patients on UC-PC recommendation margin: potential PC (UC/ED/UC or ED) utilization rates of the two counterfactuals run in parallel, and the gap in potential utilization of any care decreases over time.

Focusing on the levels of potential utilization rates, we find that patient healthcare take-up is generally low. On ED-UC recommendation margin, the potential ED utilization rate is only 43.6%,[20] and the potential utilization of any care (PC, UC, or ED) is only 59.8% within three days when the patient is recommended ED (Fig. 2). 28 days after the call, the potential utilization of any care is around 80%, which means that one in five patients does not receive any professional care for four weeks after the call, even though they are recommended the most intensive form of care (ED). Similarly, on UC-PC recommendation margin, over one in five patients do not receive any professional care for four weeks after the call, even when recommended UC (Fig. 3). With this high level of patient non-utilization as a benchmark, our TSLS estimates suggest that recommending more intensive care over less intensive care has a substantial impact on patient choices.

Methodologically, we note that, under the IV condition, the average potential responses for binary outcomes must be bounded between 0 and 1 (Chan et al., 2022; Kitagawa, 2015).[21] The estimated potential outcomes in Figs. 2–4 are consistent with this implication, providing stronger support for the validity of our IV strategy.

[17] Recall that, in addition to baseline controls X_i^0 (call center-by-call time, call center-by-algorithm disposition, and patient county), our preferred specification also controls for hold-out controls H_i (age, sex, marital status, race/ethnicity, period of service, prior healthcare utilization, and prior diagnoses) for statistical precision as shown in Eq. (3).

[18] In results not shown, we complete all robustness exercises on the UC-PC and PC-Home margins, again with little change in the size or precision of the estimates.

[19] We estimate $E\left[Y_i(r) \mid i \in \text{Complier}, X_i^0\right]$ for $r \in \{ED, UC\}$, following Abadie (2002). For $r = ED$, we regress $Y_i \cdot R_i^{ED}$ on R_i^{ED}, where the right-hand side R_i^{ED} is instrumented by the leave-out measure Z_i^{ED}. The TSLS estimate for the right-hand side R_i^{ED} identifies the average potential outcomes of compliers. For $r = UC$, we run the same regression after replacing R_i^{ED} with $1 - R_i^{ED}$. See Appendix Section A7 for details.

[20] In contrast, Tran et al. (2017) reports that 68.6% of the patients recommended ED show up ED in an Australian telephone triage system.

[21] Chan et al. (2022) uses a bound between −1 and 0 for this test, as they run a TSLS regression of $outcome \cdot (1 - treatment)$ on $treatment$. In contrast, our potential outcomes are estimated by running a TSLS regression of $outcome \cdot treatment$ on $treatment$.

5.5. Average characteristics of compliers

While the LATEs represent causal effects among compliers to whom different nurses would give a different recommendation, the underlying complier population differs between the three margins (ED-UC, UC-PC, PC-Home) and from the overall sample in observable and unobservable ways. To better interpret our TSLS estimates, we profile observable characteristics of our complier population on each margin.[22]

Table 5 contrasts the average characteristics of the compliers on the three margins and those of the overall sample. The estimated averages of prior diagnoses (Elixhauser comorbidity counts, max = 31) are 3.003 (ED-UC), 2.742 (UC-PC), 2.628 (PC-Home) (overall mean = 2.901). The estimates of complier age are 63.043 (ED-UC), 60.772 (UC-PC), and 60.477 (PC-Home) (overall mean = 63.054). Similarly, Appendix Tables B20–B23 further examine each of the 31 prior diagnoses. Again, we find that, overall, the prevalence of comorbidity decreases along the recommendation margins: highest on the ED–UC margin, lower on the UC–PC margin, and lowest on the PC–Home margin.

As described in Section 4.2, we derive the margin-specific causal interpretation of the TSLS estimates under the assumption that nurses select a recommendation based on patient acuity (ED for the highest acuity, UC for upper middle, PC for lower middle, and Home for the lowest), forming the high (ED-UC), middle (UC-PC), and low (PC-Home) acuity margins. The observed patterns of the average complier age and prior diagnoses – important correlates of patient acuity – are consistent with this framework.

6. Conclusion

This paper studies how triage recommendations affect patient utilization, healthcare costs, and health outcomes, exploiting unique data from the telephone triage system for US veterans. Extending the recent methodological developments in the examiners design literature into a healthcare setting, we estimate margin-specific causal effects of nurse recommendations on three acuity margins (ED-UC, UC-PC, PC-Home) using a focal recommendation propensity as an instrument for the focal recommendation, while controlling for non-focal recommendation propensities.

Our results have three key takeaways. First, nurse recommendations are highly influential in determining and shifting patient choices upstream. Recommending more acute care over less acute care (e.g., ED over UC) substantially influences patient utilization on the corresponding margins, while also modestly influencing patient use of professional care on the extensive margin. Second, these induced shifts in care settings impact healthcare costs, and while ED visits reduce the probability that a patient needs to call the triage line again soon after, we do not find any evidence that these shifts result in survival gain. We find that recommending UC results in lower costs on both high (ED-UC) and middle (UC-PC) acuity margins, suggesting that for marginal patients, urgent care is cost-saving in addressing acute healthcare needs. Third, we find the high prevalence of non-utilization of professional care among the callers. With the high non-utilization as a benchmark, our results suggest that adjusting triage recommendations can play an important role in encouraging patients to see a healthcare provider.

As for real-world costs of human variability, the potential gain from eliminating nurse variation is likely modest. For concreteness, we perform a back-of-the-envelope calculation when re-assigning a patient on ED-UC margin to a nurse with a standard deviation lower ED tendency (SD = 7.3 p.p.). This patient is more likely to be recommended UC over ED by 6.3 percentage points (0.86 (FS) $\times$ 0.073). This change reduces total costs by \$25.5 (\$404 $\times$ 0.063) (1.4%). The implied cost reduction is likely a modest sum, especially compared to the potential costs of aligning nurse behavior.

There are some issues that remain for future research. First, our study does not compare patients who go through telephone triage and those who do not. To have a complete picture of telephone triage, future research will need to examine how telephone triage and the lack thereof influence patient outcomes. Second, our design does not allow us to study patients to whom the nurses do not vary in their recommendations. These subsets include, for example, patients with very high acuity to whom all nurses would unanimously recommend ED care. Triage and subsequent care likely play an important role in determining outcomes of these patients. Alternative research designs and identifying variations are required to study these patients. Lastly, since triage nurse assignment does not fully determine care in our setting, our design does not allow us to directly estimate the effect of receiving actual care (e.g., ED care). Future research may enrich the examiners design literature by exploring identification conditions for settings where the first decision-maker affects the second decision-maker's choice.

CRediT authorship contribution statement

Liam Rose: Writing – original draft, Data curation, Conceptualization. **Ken Suzuki:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Linda Diem Tran:** Data curation, Conceptualization. **Anita Vashi:** Funding acquisition, Conceptualization.

[22] Following Abadie (2002), we estimate $E\left[h_{ki} \mid i \in \text{Complier}, X_i^0\right]$ for some hold-out characteristic h_{ki} by regressing $h_{ki} \cdot R_i^{ED}$ on R_i^{ED} with the right-hand side R_i^{ED} being instrumented by the leave-out measure Z_i^{ED}. See Appendix Section A8 for details.

Declaration of competing interest

Liam Rose declares that over the past three years, the author received research grants from the Department of Veterans Affairs, the National Institutes of Health, and the American Cancer Society.

No party had the right to review the submitted paper.

Ken Suzuki declares that over the past three years, the author received research grants from the Department of Veterans Affairs and the University of California, Santa Cruz.

No party had the right to review the submitted paper.

Linda Diem Tran declares that over the past three years, the author received research grants from the Department of Veterans Affairs.

No party had the right to review the submitted paper.

Anita Vashi declares that over the past three years, the author received research grants from the Department of Veterans Affairs.

No party had the right to review the submitted paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhealeco.2026.103149.

Data availability

This paper analyzes administrative health records from the US Department of Veterans Affairs. This data cannot be made publicly available due to the sensitive nature of individual-level health records.

References

Abadie, Alberto, 2002. Bootstrap tests for distributional treatment effects in instrumental variable models. J. Amer. Statist. Assoc. 97 (457), 284–292.

Abadie, Alberto, 2003. Semiparametric instrumental variable estimation of treatment response models. J. Econometrics 113 (2), 231–263.

Abualenain, Jameel, Frohna, William J., Shesser, Robert, Ding, Ru, Smith, Mark, Pines, Jesse M., 2013. Emergency department physician-level and hospital-level variation in admission rates. Ann. Emerg. Med. 61 (6), 638–643.

Alexander, Diane, Currie, Janet, Schnell, Molly, 2019. Check up before you check out: Retail clinics and emergency room use. J. Public Econ. 178, 104050.

Allen, Lindsay, Cummings, Janet R., Hockenberry, Jason M., 2021. The impact of urgent care centers on nonemergent emergency department visits. Health Serv. Res. 56 (4), 721–730.

Arteaga, Carolina, 2023. Parental incarceration and children's educational attainment. Rev. Econ. Stat. 105 (6), 1394–1410.

Bhalotra, Sonia R., Daysal, N. Meltem, Trandafir, Mircea, 2025. Antidepressant Treatment in Childhood. Technical Report, JSTOR.

Chan, David C., Card, David, Taylor, Lowell, 2023. Is there a VA advantage? evidence from dually eligible veterans. Am. Econ. Rev. 113 (11), 3003–3043.

Chan, David C., Gentzkow, Matthew, Yu, Chuan, 2022. Selection with variation in diagnostic skill: Evidence from radiologists. Q. J. Econ. 137 (2), 729–783.

Chan, David C., Gruber, Jonathan, 2020. Provider discretion and variation in resource allocation: The case of triage decisions. AEA Pap. Proc. 110, 279–283.

Chyn, Eric, Frandsen, Brigham, Leslie, Emily C., 2024. Examiner and Judge Designs in Economics: A Practitioner's Guide. National Bureau of Economic Research Working Paper Series, No. 32348.

Costa-Ramon, Ana, Daysal, Meltem, Ana, Rodríguez-González, 2023. The Oral Contraceptive Pill and Adolescents' Mental Health. Technical Report, CEPR Discussion Papers.

Coussens, Stephen, Ly, Dan P., 2025. Variation in emergency department physician admitting practices and subsequent mortality. JAMA Intern. Med. 185 (2), 153–160.

Currie, Janet, Karpova, Anastasia, Zeltzer, Dan, 2023. Do urgent care centers reduce medicare spending? J. Health Econ. 89, 102753.

Currie, Janet, Zwiers, Esmée, 2025. Medication of postpartum depression and maternal outcomes: Evidence from geographic variation in dutch prescribing. J. Hum. Resour. 60 (4), 1093–1125.

Dalsgaard, Søren, Nielsen, Helena Skyt, Simonsen, Marianne, 2014. Consequences of ADHD medication use for children's outcomes. J. Health Econ. 37, 137–151.

Eichmeyer, Sarah, Zhang, Jonathan, 2022. Pathways into opioid dependence: Evidence from practice variation in emergency departments. Am. Econ. J.: Appl. Econ. 14 (4), 271–300.

Eichmeyer, Sarah, Zhang, Jonathan, 2023. Primary care providers' influence on opioid use and its adverse consequences. J. Public Econ. 217, 104784.

Ferro, Simone, Serra, Chiara, 2025. Triage at shift changes and distortions in the perception and treatment of emergency patients. J. Health Econ. 99, 102944.

Flores-Mateo, Gemma, Violan-Fors, Concepción, Carrillo-Santisteve, Paloma, Peiró, Salvador, Argimon, Josep-Maria, 2012. Effectiveness of organizational interventions to reduce emergency department utilization: A systematic review. PLOS ONE 7 (5), e35903, Publisher: Public Library of Science.

Frandsen, Brigham, Lefgren, Lars, Leslie, Emily, 2023. Judging judge fixed effects. Am. Econ. Rev. 113 (1), 253–277.

Humphries, John Eric, Ouss, Aurelie, Stavreva, Kamelia, Stevenson, Megan T., van Dijk, Winnie, 2024. Conviction, incarceration, and recidivism: Understanding the revolving door.

Islam, Farah, Sabbe, Marc, Heeren, Pieter, Milisen, Koen, 2021. Consistency of decision support software-integrated telephone triage and associated factors: a systematic review. BMC Med. Inform. Decis. Mak. 21 (1), 107.

Kamat, Vishal, Norris, Samuel, Pecenco, Matthew, 2024. Conviction, incarceration, and policy effects in the criminal justice system. SSRN Electron. J.

Kirkeboen, Lars J., Leuven, Edwin, Mogstad, Magne, 2016. Field of study, earnings, and self-selection*. Q. J. Econ. 131 (3), 1057–1111.

Kitagawa, Toru, 2015. A Test for instrument validity. Econometrica 83 (5), 2043–2063, Publisher: [Wiley, The Econometric Society].

Kline, Patrick, Walters, Christopher R., 2016. Evaluating public programs with close substitutes: The case of head start*. Q. J. Econ. 131 (4), 1795–1848.

Molitor, David, 2018. The evolution of physician practice styles: Evidence from cardiologist migration. Am. Econ. J.: Econ. Policy 10 (1), 326–356.

Morgan, Katherine Majzoub, Karadakic, René, Barnett, Michael L, 2026. Primary care clinicians available for new patient visits. JAMA Intern. Med.

Mueller-Smith, Michael, 2015. The criminal and labor market impacts of incarceration. p. 59.

Pinto, Rodrigo, 2022. Beyond intention-to-treat: Using the incentives of moving to opportunity to identify neighborhood effects.

Raven, Maria C., Kushel, Margot, Ko, Michelle J., Penko, Joanne, Bindman, Andrew B., 2016. The effectiveness of emergency department visit reduction programs: A systematic review. Ann. Emerg. Med. 68 (4), 467–483.e15.

Retchin, Sheldon M., Garland, Sheryl L., Anum, Emmanuel A., 2009. The transfer of uninsured patients from academic to community primary care settings. Am. J. Manag. Care 15 (4), 245–252.

Sadowski, Laura S., Kee, Romina A., VanderWeele, Tyler J., Buchanan, David, 2009. Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: A randomized trial. JAMA 301 (17), 1771–1778.

Sexton, Vanashree, Dale, Jeremy, Bryce, Carol, Barry, James, Sellers, Elizabeth, Atherton, Helen, 2022. Service use, clinical outcomes and user experience associated with urgent care services that use telephone-based digital triage: a systematic review. BMJ Open 12 (1), e051569.

Sigstad, Henrik, 2023. Monotonicity Among Judges: Evidence from Judicial Panels and Consequences for Judge IV Designs. Working Paper.

Silver, David, Zhang, Jonathan, 2022. Invisible Wounds: Health and Well-Being Impacts of Mental Disorder Disability Compensation on Veterans. National Bureau of Economic Research Working Paper Series, No. 29877.

Soril, Lesley J.J., Leggett, Laura E., Lorenzetti, Diane L., Noseworthy, Tom W., Clement, Fiona M., 2015. Reducing frequent visits to the emergency department: A systematic review of interventions. PLOS ONE 10 (4), e0123660, Publisher: Public Library of Science.

Tran, Duong Thuy, Gibson, Amy, Randall, Deborah, Havard, Alys, Byrne, Mary, Robinson, Maureen, Lawler, Anthony, Jorm, Louisa R., 2017. Compliance with telephone triage advice among adults aged 45 years and older: an Australian data linkage study. BMC Health Serv. Res. 17 (1), 512.

Vashi, Anita A., Urech, Tracy, Carr, Brendan, Greene, Liberty, Warsavage, Jr., Theodore, Hsia, Renee, Asch, Steven M., 2019. Identification of emergency care–sensitive conditions and characteristics of emergency department utilization. JAMA Netw. Open 2 (8), e198642.

Wang, Bill, Mehrotra, Ateev, Friedman, Ari B., 2021. Urgent care centers deter some emergency department visits but, on net, increase spending. Health Aff. 40 (4), 587–595, Publisher: Health Affairs.

Weinick, Robin M., Burns, Rachel M., Mehrotra, Ateev, 2010. Many emergency department visits could be managed at urgent care centers and retail clinics. Health Aff. 29 (9), 1630–1636, Publisher: Health Affairs.