Evaluating the Efficacy of Connected Healthcare: An Empirical Examination of Patient Engagement Systems and Their Impact on Readmission

Tan Lekwijit
Operations, Information and Decisions, The Wharton School, University of Pennsylvania, slek@wharton.upenn.edu

Christian Terwiesch
Operations, Information and Decisions, The Wharton School, University of Pennsylvania, terwiesch@wharton.upenn.edu

David A. Asch
Perelman School of Medicine, University of Pennsylvania, asch@wharton.upenn.edu

Kevin G. Volpp
Perelman School of Medicine, University of Pennsylvania, volpp70@wharton.upenn.edu

Connected healthcare is a form of health delivery that connects patients and providers through connected health devices, allowing providers to monitor patient behavior and proactively intervene before an adverse event occurs. Unlike the costs, the benefits of connected healthcare in improving patient behavior and health outcomes are usually difficult to determine. In this study, we examine the efficacy of a connected health system that aimed to reduce readmissions through improved medication adherence. Specifically, we study 1,000 patients with heart disease who received electronic pill bottles that tracked medication adherence. Patients who were non-adherent received active social support that involved different types of feedback such as text messages and calls. By integrating data on adherence, intervention, and readmission, we aim to (1) investigate the efficacy of connected healthcare in promoting medication adherence, (2) examine the relationship between medication adherence and readmission, and (3) develop a dynamic readmission risk-scoring model that considers medication adherence and use the model to better target non-adherent patients. Our findings suggest that patients are more likely to become adherent when they or their partners receive high levels of intervention that involve personalized feedback and when the intervention is escalated quickly and consistently. We also find that long-term adherence to two crucial heart medications, statins and β-blockers, is strongly associated with reduced readmission risk. Lastly, using counterfactual simulation, we apply the dynamic readmission risk-scoring model to our setting and find that, when using an intervention strategy that prioritizes high-risk patients, we obtain 10% fewer readmissions than we would obtain without considering readmission risk while using the same effort level from the patient support team.

Key words: connected healthcare; health information technology; medication adherence; readmission analytics; behavioral interventions

1. Introduction

More than ever, technology is bringing businesses and customers closer together. In the past, consumers and their banks would interact only sporadically in a retail branch. Now, online and mobile banking connect the two parties almost seamlessly, creating a delivery model known as connected banking. Similarly, automotive insurance companies used to interact with the drivers they insured only at predefined touch points, such

as policy renewals or claims events. Now, sensors and mobile apps continuously monitor driving behavior. Such data enable a connected insurance business model by providing safety feedback to drivers (and parents) as well as informing future underwriting decisions.

The domain that we study in this paper is healthcare. In the past, before the possibility of being connected technologically to the healthcare system, patients made decisions related to diet, exercise regimen, or medication adherence without the involvement of care providers (Asch et al. 2012). Now, connected health devices, such as smart pill bottles, connected scales, and wearable health trackers, are capable of transmitting clinically relevant data to information systems that can then relay this information to patients and providers. The promise of connected healthcare is that such information allows healthcare providers to monitor health-related behavior and to proactively intervene rather than waiting for an adverse event. Moreover, it has been suggested that patients adjust their behavior in response to feedback, becoming more adherent in taking their medication and more active in their lifestyle.

Active patient engagement and feedback systems have the potential to reduce the number of preventable readmissions and ultimately reduce healthcare costs (Chouvarda et al. 2015). Despite this enormous upside, however, there is limited evidence for the effectiveness of connected healthcare and little guidance on how to effectively implement it (Caulfield and Donnelly 2013). One of the crucial health behaviors that may affect the readmission risk and have been the target for improvement is medication adherence. Although the costs of implementation are usually straightforward to estimate, the benefits of engagement systems to improve medication adherence and the impact of higher adherence on readmission are oftentimes ambiguous. Moreover, it is unclear how to effectively engage with non-adherent patients in order to minimize the number of readmissions given limited resources.

The aims of our research are to (1) investigate the efficacy of connected feedback systems in promoting medication adherence among patients, (2) examine the relationship between patient medication adherence and readmission, and (3) develop a readmission risk-scoring model that takes into account medication adherence and use the model to better target non-adherent patients. Establishing how connected health systems causally affect clinical outcomes can be done only in a large-scale randomized controlled trial. Running such an intervention requires overcoming multiple challenges. First, one needs to recruit a large cohort of patients, equip them with connected devices, and train them how to use the devices. Second, for patients to benefit from connected healthcare, new delivery processes have to be designed and implemented. Third, one has to track patients in the study not only during their usage of the connected healthcare devices, but also for a long period so that the full scope of the clinical outcomes can be observed.

Our study is part of a larger research project that overcame these challenges by implementing one of the largest clinical interventions related to connected healthcare that has been conducted so far. This study, known as the HeartStrong study, enrolled 1,453 patients for a two-arm randomized clinical trial with a year-long intervention that aimed to improve medication adherence and reduce the number of readmissions

after myocardial infarction. The 1,000 patients in the intervention group received a compound intervention integrating wireless pill bottles that can electronically track openings and transmit them to the care team, lottery-based incentives, and social support that involved feedback systems such as automated messaging, manual messaging, and phone calls, delivered to either the patients themselves or their enlisted partner at increasing levels of escalation. Patients were then tracked for 12 months, capturing any future hospital readmissions.

The results of the study were published in one of the leading medical journals (Volpp et al. 2017) and did not show a statistically significant difference in readmission between the control group and the intervention group. The lack of statistically significant difference begs the question of why patients who were better connected to the healthcare system did not benefit in the form of better outcomes. Did the various reminder and feedback interventions improve medication adherence? Did higher rates of adherence reduce the likelihood of readmission? Did we sufficiently allocate our capacity to the patients who were at high risk of readmission? Without a more micro-level theory, one can only speculate.

The contribution of this study is to present such a micro-level operational model. To evaluate the effectiveness of the connected health system, the HeartStrong study focused on the variation between the control group and the intervention group and did not use any data related to behavioral interventions and daily medication adherence. In contrast, the present study exploits variation within the intervention group, leveraging the micro-level data capturing patient behavior and provider actions at the patient-day level. At this micro level, hundreds of observations per patient (did patient i receive an intervention j on day t?) were analyzed and linked with adherence behavior (did patient i take medication k on day t?) and clinical outcomes (was patient i readmitted on day t?). This allowed us to establish the following novel contributions:

1. *Showing the impact of escalation strategies on medication adherence*: By studying the effect of different levels and dynamics of intervention escalation, we show that patients are significantly more likely to become adherent when (a) they or their partners receive high levels of intervention that involve personalized feedback like phone calls and manual messages and (b) the intervention is escalated quickly and consistently. Our results imply that, in order to effectively make a previously non-adherent patient adherent again, one should start calling the patient as soon as he or she becomes non-adherent. Receiving a personal phone call immediately after the first day of non-adherence more than doubles the probability of becoming adherent again. Then, if non-adherence continues, one should escalate to manual messages and calls to the patient's partner successively on the following days.
2. *Showing the impact of medication adherence on readmission*: We explore how consistency in medication use in the past affects the readmission probability in the present. Specifically, we show that, for patients with cardiovascular disease, long-term adherence to statins and β-blockers is associated with a 51% reduction in the odds of being readmitted on any given day, thus better medication adherence is strongly associated with reduced risk of readmission.

3. *Developing a dynamic readmission risk-scoring model and using it to better target non-adherent patients*: Using multi-layer perceptron, we develop a dynamic readmission risk-scoring model that includes patient-day level medication adherence as predictors. We evaluate the model primarily using the area under the ROC curve and find that (a) the model outperforms a baseline model that does not include medication adherence and (b) the model outperforms other models developed using different machine learning methods. Moreover, using counterfactual analysis, we apply the dynamic readmission risk-scoring model to our setting and show that, when using an intervention strategy that prioritizes patient-days with highest readmission risks, we obtain 10% fewer readmissions than we would obtain without considering readmission risk while maintaining the same level of effort.

Our connected healthcare setting holds great significance because reducing readmission rates of patients with heart failure is a national priority (Bradley et al. 2013). Nearly 20% of Medicare beneficiaries discharged from the hospital are readmitted within 30 days, and these readmissions have been estimated to cost the country more than $24 billion (Agency for Healthcare Research and Quality 2014). Since a significant part of the readmission risk can be attributed to patients' lack of adherence to their prescribed medications, real-time adherence monitoring and reminder systems hold particular promise.

The rest of the paper is organized as follows. We provide a brief summary of related papers in the literature in Section 2. In Section 3, we introduce our study setting. We present the study on the impact of escalation strategies on adherence behavior and its estimation results in Section 4. The study on the impact of medication adherence on readmission is presented in Section 5. We develop a dynamic readmission-risk scoring model and apply it to our setting in Section 6. Section 7 provides concluding remarks as well as discussions for future research.

2. Literature Review

The adoption of remote monitoring and feedback has long been prevalent outside of the healthcare industry. Researchers in operations management have been interested in examining remote monitoring, diagnosis, and feedback in manufacturing and other service settings. In manufacturing, the focus is usually on monitoring people, processes, and machines. Information technology is often adopted to reduce the cost of capturing, storing, and transmitting data to all members of the supply chain as well as to enable useful feedback functions such as fault notification, remote counseling, and real-time online help and process intervention. In service industries, remote monitoring and feedback became popular thanks to the call for digitalization in services (Tan and Netessine 2019) and the rapid growth of teleservices, which allow customers to receive service virtually at their convenience and, at the same time, allow behavioral data to be collected and shared with service providers. An increasing number of studies explore the prospects of utilizing such technologies in popular service domains such as retail and banking, through the use of e-commerce and online banking (e.g., Campbell and Frei 2010, Moon et al. 2018). In these service settings,

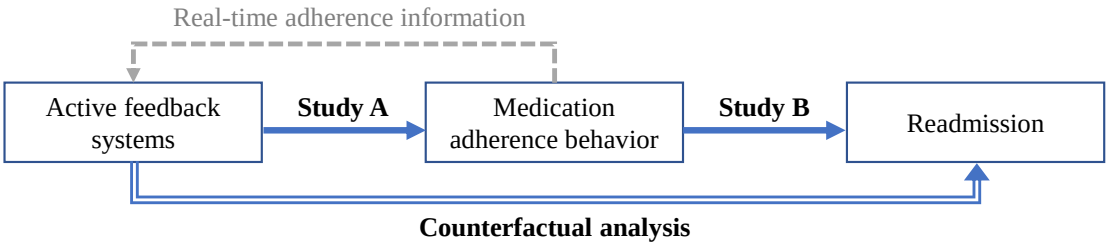
real-time information sharing and customer interaction can benefit service providers by helping them target customers with specific price promotions, detect fraudulent activities, improve customer retention, and increase market share. This leads to "connected strategies" that allow firms to build continuous relationships with customers by having frequent, low-friction interactions with them and to address their needs as or even before those needs arise (Siggelkow and Terwiesch 2019). Apart from their use with customers, remote monitoring and feedback are also increasingly utilized to oversee internal service operations such as monitoring employee theft and productivity (e.g., Pierce et al. 2015).

More recently, the applications of remote monitoring and feedback have expanded into the healthcare community. For healthcare workers, remote electronic monitoring can be put in place to help increase workers' behavioral compliance, such as hand hygiene (Staats et al. 2017). For patients, telehealth practices have been increasingly adopted to facilitate efficient management of health and wellness, allowing health personnel to proactively connect with patients through health IT and provide patients with real-time health and behavioral assessment, diagnosis, interventions, and consultation (Kvedar et al. 2014). As more researchers and practitioners are interested in learning how health IT may transform healthcare practices, we see a growing body of operations management literature that explores the impact of various long-standing forms of health IT and telehealth, such as home monitoring and e-visits (e.g., Rajan et al. 2019, Bavafa et al. 2018, Angst et al. 2011, Devaraj et al. 2013).

In our study, we focus on connected healthcare that provides remote monitoring and feedback for patients with heart disease. A number of past studies investigate the benefits of connected health, or more broadly, telemonitoring, in patients with chronic conditions (e.g., Watson et al. 2009, Trappenburg et al. 2008), with many focusing on heart conditions (e.g., Chaudhry et al. 2010, Cleland et al. 2005, Maeng et al. 2014). Although past studies, which primarily rely on small- to medium-scale randomized and non-randomized control trials, suggest that telemonitoring may be an effective strategy for disease management in heart failure patients, the evidence base is inconclusive and quite limited (Chaudhry et al. 2007). Moreover, there exists little guidance on how to effectively implement such a system.

Our study aims to (1) examine the efficacy of connected health systems in improving medication adherence, (2) examine the relationship between medication adherence and readmission, and (3) develop a dynamic readmission risk-scoring model that takes medication adherence as input and use it to better target non-adherent patients. As shown in Figure 1, we address our first two aims in two studies, Study A and Study B. To address the third aim, we use machine learning to develop a dynamic risk-scoring model and investigate the benefit of utilizing the model in our intervention delivery through counterfactual simulations, which link intervention to readmission by treating medication adherence as a mediator. Our work joins an increasing number of studies in operations that focus on remote monitoring and compliance (e.g., Staats et al. 2017, Jonasson et al. 2020), applications of data analytics in healthcare (e.g., Wang et al. 2019),

Figure 1 Analysis Outline



chronic disease management (e.g., Jonasson et al. 2017), and behavioral healthcare analysis (e.g., Ibanez et al. 2018), particularly analysis of patient behavior (e.g., Liu et al. 2018).

In the first study (Study A), we investigate the effectiveness of different escalation strategies in promoting medication adherence behavior. The lack of medication adherence has been a serious problem in managing chronic diseases around the world. Numerous studies have shown that patients with chronic illnesses adhere to their prescribed medications only 50% to 60% of the time, and the total direct and indirect cost estimates for non-adherence range from $100 billion to $300 billion each year (Bosworth et al. 2011). Until now, traditional behavioral interventions targeting medication adherence have produced only modest success (e.g., Ho et al. 2009, Nieuwlaat et al. 2014). Furthermore, the evidence regarding the effectiveness of health IT interventions to improve adherence is fairly thin (Bosworth et al. 2011). Bosworth et al. (2011) has found that almost all of the interventions that were effective were complex, including combinations of convenient care, information, reminders, self-monitoring, reinforcement, counseling, family therapy, and crisis intervention. The finding is in line with Ho et al. (2009), which concludes that multimodal interventions have been more successful than unimodal interventions, which rely on methods like reducing the number of daily doses of medications or packaging medications into special containers. These findings consistently suggest that connected healthcare, which can deliver multimodal interventions at reasonable cost, is a promising care model.

To the best of our knowledge, no one has attempted to study how patients' medication adherence behavior changes in response to different escalation strategies. Prior studies that study the impact of adherence intervention on medication adherence primarily rely on randomized controlled trials and before-and-after comparisons. They work with macro-level intervention information (e.g., always receiving reminders vs. never receiving reminders) and approximate aggregate adherence data (e.g., self-reported adherence or prescription fill rates in the past three months). These studies do not have the granularity of data that we have available and therefore do not examine directly how patients respond to different escalation levels and dynamics. In contrast, we conduct our analysis at the patient-day level by utilizing micro-level intervention and adherence data, which allow us to examine the relationship between different escalation strategies and patients' immediate changes in behavior (or lack thereof). Micro-level data recorded over time are

commonly used in healthcare operations studies, especially to analyze workers' behavior and performance (e.g., Gurvich et al. 2019, Berry Jaeker and Tucker 2017).

In the second study (Study B), we examine the impact of adherence to two crucial medications, statins (a class of lipid-lowering medications; also known as HMG-CoA reductase inhibitors) and β-blockers (a class of medications that are particularly used to manage cardiac arrhythmias; also written as beta-blocker) on the risk of readmission. We focus on readmission as the outcome variable for two main reasons (Kansagara et al. 2013). First, the ability to assess the readmission risk helps physicians target the delivery of care, especially resource-intensive interventions, to patients who are at highest risk for readmission. Second, readmission rates are often used as a care quality measure (e.g., Kim et al. 2015, Chan et al. 2019) and a quality metric for healthcare providers. Being able to identify drivers of readmission is important for hospitals. Since 2012, the Centers for Medicare and Medicaid Services (CMS) has publicly reported readmission rates and planned to lower reimbursement to hospitals with high risk-standardized readmission rates.

The relationship between adherence and health outcomes is recognized but understudied. DiMatteo et al. (2002) conducted a meta-analysis and proposed that the relationship was under-investigated because the effect of adherence on outcomes was often taken for granted. Nevertheless, several studies have attempted to determine the impact of medication adherence on health outcomes such as rehospitalization and mortality for patients with chronic diseases such as coronary heart disease, diabetes, and AIDS (e.g., McDermott et al. 1997, Yu et al. 2010, Han et al. 2014). However, the difficulties of coming up with effective experimental study designs and having an accurate measure of adherence have made it difficult for researchers to establish a clear causal relationship between medication adherence and outcomes. Most previous studies rely on traditional measures of adherence such as self-reports, physician reports, and prescription fill rates. They cannot accurately measure medication adherence (Lam and Fresco 2015) and obtain medication adherence information directly prior to each readmission. In contrast, our study utilizes medication adherence data from electronic pill bottles. This allows us to analyze the probability of readmission at the patient-day level using reliable adherence information leading up to each day using an econometric approach, as opposed to traditional cross-sectional comparisons.

Finally, in our third analysis, we develop a dynamic readmission risk-scoring model that includes medication adherence as predictors and use counterfactual simulations to examine the benefit of using the model in our intervention delivery. Instead of alerting every patient at pre-specified moments of non-adherence, our new strategy directly targets the patients (and the patient-days) with highest readmission risks. Many researchers have explored various ways data analytics can improve healthcare delivery (e.g., Raghupathi and Raghupathi 2014, Wang et al. 2019). However, to the best of our knowledge, none of them has studied the application of predictive analytics in connected health delivery. Moreover, although many studies attempt to develop readmission prediction models (e.g., Min et al. 2019, Shulan et al. 2013), they mostly aim to predict the probability of readmission within a fixed period after hospital discharge (e.g.,

probability of readmission within 30 days). Unlike our study, most prior studies do not have post-discharge patient-day level covariates, barring them from obtaining patient-day level readmission risk.

3. Study Setting

This study builds on a national two-arm randomized controlled trial program we conducted from 2013 to 2016. We collected data from 1,000 patients who enrolled in the program and received a 12-month connected healthcare intervention that aimed to reduce repeated cardiovascular events through improved medication adherence. Our original study was one of the first studies to deploy connected health devices on a large scale. The study aimed to improve upon earlier intensive case management efforts, which were mostly unwieldy and expensive, by using a simpler, cheaper, and more scalable approach. By leveraging more connected technology, we were able to deploy a novel, proactive model of chronic disease management. We also created new workforce roles and shifted some intensive case management roles from physicians to non-physician providers, which can benefit both physicians and the hospital (Powell et al. 2012).

Our study was approved by the Institutional Review Board of the University of Pennsylvania. Prior to the study, patients were recruited by University of Pennsylvania research staff from 2013 through 2015 and observed for one year. Eligible patients were:

(a) 18 to 80 years old.
(b) Admitted as hospital inpatients for one to 180 days.
(c) Discharged home with a primary diagnosis code of acute myocardial infarction (AMI).
(d) Not suffering from dementia.
(e) Not enrolled in other research studies incorporating wireless pill bottles.

Patients could enroll up to 60 days after discharge. To recruit patients, we initially contacted them through letters and phone calls. Of 19,678 potentially eligible patients contacted, 18,169 declined, could not be reached, or were ineligible upon further inspection, leaving an enrolled sample of 1,509. The most common reasons cited by the patients who declined were: (a) not being interested, (b) not wanting to change the current system, and (c) having privacy concerns. All enrolled patients received $25 for participation. Patients in our sample were insured with five large US insurers or with Medicare fee-for-service at the University of Pennsylvania Health System.

We randomized patients in a 2:1 ratio of intervention:usual care using permuted block randomization stratified by insurance provider to balance the allocation across provider groups. After excluding patients who withdrew or did not have post-enrollment medical claims, 975 patients from the intervention group remain in our final patient cohort. Patients in the intervention group received an additional $25 for activating wireless pill bottles (Vitality GlowCaps), which electronically monitored bottle openings with a small remote device that plugged into a wall outlet and transmitted cellular signals to the care team, with no home wireless network, computer, or special setup required. Adherence information transmitted from those pill

bottles was recorded on an electronic platform that we primarily utilized to facilitate real-time monitoring and feedback. Our connected healthcare platform, which was developed with funding from the National Institutes of Health, is a flexible and secure web-based infrastructure that consists of a portal that can be linked to various peripheral connected health devices, such as scales, glucometer, and pill bottles. The platform is capable of automating the delivery of feedback and communicating back to patients using email, text messaging, and interactive voice recording.

Patients in the intervention group were assigned an engagement advisor (EA) for the duration of their participation. The EAs assisted patients with pill bottle setup and troubleshooting, monitored the patient's daily medication adherence on the electronic platform, provided manual forms of feedback, and served as a resource for patients struggling to stay adherent. All intervention patients were asked to enlist a potential support partner, usually a friend or family member. The role of the support partner was to receive information about the patient's adherence through the platform and EAs, and provide support and encouragement using their preferred format and content. Support partners also had an account on the electronic platform, allowing them to automatically receive alerts through email, text message, or automated phone call. We gave intervention patients access to social work resources as well as five main forms of adherence feedback with increasing degrees of intensity: (1) automated message to patient, (2) automated message to partner, (3) phone call to patient, (4) manual message to patient, and (5) phone call to partner. We discuss different forms of intervention in further detail in Section 4.2 in conjunction with Study A. Patients in the control group received usual care for the duration of the study and had no further contact with EAs or study staff. The results of our randomized controlled analysis are described in Volpp et al. (2017). We did not find statistically significant differences between study arms in time-to-first-rehospitalization for a vascular event or death, or total number of repeated hospitalizations

To investigate further if and how the connected heath systems worked, this paper focuses on the patients who received the intervention. These intervention patients had within-person variation in their adherence behavior and variation in how the intervention was escalated. The data we use for our micro-level analysis contains both operational and patient-level information. Operational information includes daily data for each patient indicating whether:

(a) Each of the five forms of intervention was delivered.

(b) The patient opened each of his or her pill bottles.

(c) The patient was readmitted.

We also have information on the time between patients' initial discharge and study enrollment and on whether and when he or she was readmitted before enrollment. In addition to operational information, our data set contains information about patient characteristics such as age, gender, Medicare enrollment, Patient Health Questionnaire-2 (PHQ-2) score, and baseline Elixhauser comorbidity score. The PHQ-2, which is used to screen depression, inquires about the frequency of depressed mood and anhedonia over the past

two weeks (McManus et al. 2005, Kroenke et al. 2003). A PHQ-2 score can range from 0 to 6, with higher scores translating to higher chances of having depression. The second score, Elixhauser comorbidity score, is evaluated using up to 12 months of pre-enrollment data based on 31 individual conditions identified from diagnoses in hospital and physician data and can take positive or negative values (van Walraven et al. 2009). These scores were assigned once and not updated during the study period. Table 1 provides the summary statistics of patient demographics for all patients in our cohort.

Table 1 Summary Statistic for Patient Demographics

Patient characteristic	Mean	SD	Min	Max
Age	61.26	10.41	23	80
Female	0.34	0.47	0	1
Medicare	0.44	0.50	0	1
PHQ-2 depression score	1.26	1.63	0	6
Baseline Elixhauser score	6.46	10.03	-14	44
Time from discharge to enrollment (days)	40.52	12.36	8	60

We describe the data in more detail in Table 10 in the Appendix. The data is also made available to supplement the review process and to support further studies.

4. Study A—Impact of Remote Monitoring and Feedback Systems on Medication Adherence Behavior

We begin by examining the efficacy of the feedback systems. Specifically, we want to see how effective different escalation strategies are in turning non-adherence into adherence. By answering this question, we can develop insights into how feedback systems play a role in improving patient behavior, and how best to engage with non-adherent patients in order to make them adherent again.

4.1. Data and Statistics

The data used in this analysis contains a series of non-adherence sequences for each patient. Each non-adherence sequence begins on the first day without pill bottle use and lasts until the first day the patient takes all of his or her medications again. One patient could have multiple non-adherence sequences. Each entry in our data contains daily information on whether the patient opened all the pill bottles and whether he or she received each type of intervention. On average, we observe each patient for 318 days, including both adherence and non-adherence days. Our observation period is slightly shorter than the intervention period because the wireless pill bottle activation was occasionally delayed, causing a lag time between the start of the intervention period and when the bottles were functional.

Of 975 patients, approximately 3% were always adherent and are excluded from our analysis and 2% of the patients had irregularly long periods of non-adherence. For these patients, the average number of non-adherence days within each non-adherence sequence is greater than 15, and the maximum number of

non-adherence days within a sequence ranges from 68 to 334 days. We exclude these patients from our main analysis because they may have intentionally not taken some or all of their medications, or their pill bottles were not working properly. After those patients are excluded, our data contain 930 patients who exhibit reasonable variability in their adherence pattern. Table 2 provides adherence statistics of the remaining patients and all patients. In our final cohort, on average, each patient has 15.57 non-adherence sequences, each lasting a mean of 2.78 days.

Table 2 Summary of Medication Adherence Behavior

	Focus Cohort (N = 930)				All patients (N = 975)			
Adherence Characteristic	mean	std	min	max	mean	std	min	max
Number of non-adherence sequences per person	15.57	12.39	1	73	14.94	12.77	0	73
Number of non-adherence days within a sequence	2.78	3.05	1	122	3.86	6.10	1	334

4.2. Escalation of Feedback

We aimed to remind patients to take medications at increasing levels of escalation. Lower-level interventions were delivered automatically through our electronic platform while higher-level interventions were delivered manually by the engagement advisors who closely monitored adherence behavior. We usually provided lower-level interventions before we escalated to a higher-level intervention and stopped delivering interventions once the patient became adherent again. We illustrate the nature of escalation in Figure 2, which shows the distribution of each type of intervention over different days within a non-adherence sequence. Each mark on the vertical axis indicates the t^{th} day of non-adherence. Each of the smaller plots corresponds to each form of intervention and is displayed along with its total number of occurrences at the bottom. Each plot illustrates the percent distribution of each form of intervention over day of non-adherence and is compared across all forms of intervention, capturing the nature of escalation. For example, we delivered 122 manual messages to patients in total, 47% of which occurred on the sixth day without pill bottle use. In contrast, we initiated automated messages to patients much earlier in the non-adherence period, oftentimes in the first few days the non-adherence. Based on the escalation pattern, we define five levels of escalation corresponding to five main forms of intervention, which are described below in increasing order of escalation.

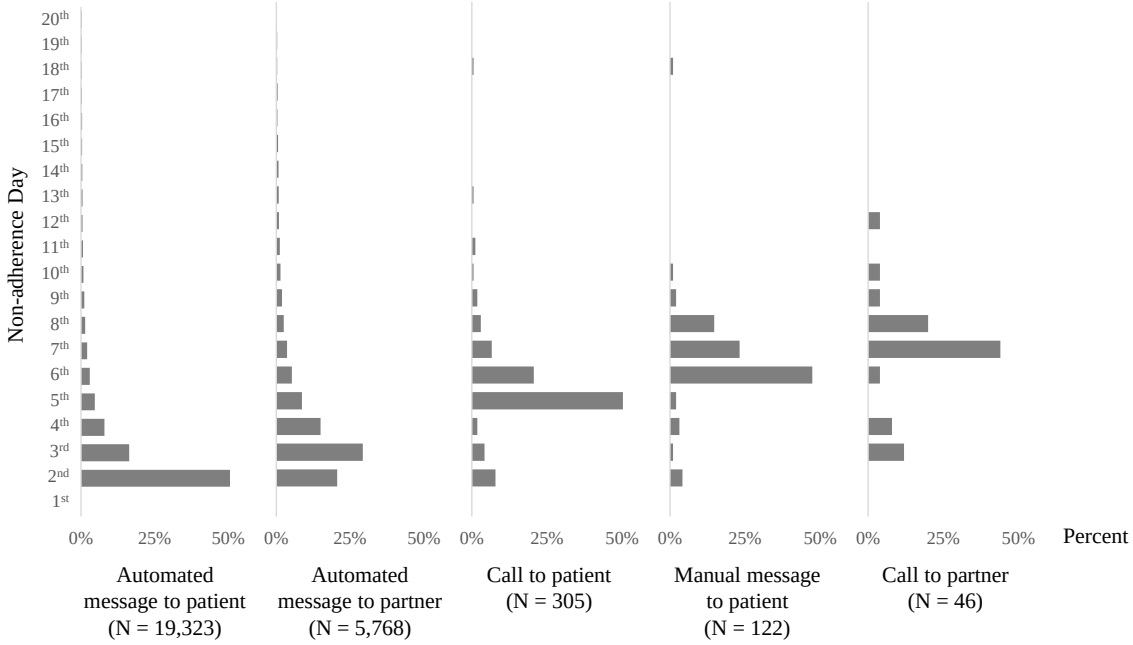
1. *Escalation level 1 — Automated message to patient*: In the lowest level of escalation, we deliver a first automated message to tell the patient that he did not take his medications. These non-personalized messages were mostly triggered via the electronic platform in the first two days of non-adherence and were usually delivered through text message or email. We usually sent automated non-adherence messages regularly until the patient became adherent again. In total, 19,323 such messages were sent. These messages contained a standard text that wrote: "You did not take all your meds yesterday."

2. *Escalation level 2 — Automated message to partner*: In the next level of escalation, we send an automated message to the patient's partner. Most non-personalized automated messages to support partners were initially triggered via the platform in the first four days of non-adherence and mostly continued until non-adherence ended. Like other platform-triggered messages, they were mainly delivered through text message and email. In total, 5,768 automated partner messages were sent. The standard message wrote: "Your HeartStrong partner has missed some heart medications for the past few days and may need your support."
3. *Escalation level 3 — Phone call to patient*: In the third level of escalation, the EAs make the first phone call to the patient after non-adherence starts. Telephone calls to patients were mostly made after the fifth day of non-adherence. The EAs directly made these calls to tell patients about their non-adherence and ask them about the reasons for why the pill bottles were not opened. In total, the EAs made 305 calls that successfully reached the patients.
4. *Escalation level 4 — Manual message to patient*: With the next escalation level, the EAs sends the first manual message to the patient after non-adherence starts. Manual messages were tailor-written and sent to the patients via the electronic platform, informing them of their non-adherence and asking them to contact the care team. The EAs mostly sent these messages after the sixth day without pill bottle use. In total, 122 manual messages were sent. The content of these messages varied; for example, a manual message may write: "Our system shows that you have not opened your GlowCaps for [number of non-adherence days] days. We have been unsuccessful in [action] so we will be contacting [person we would contact, e.g., support partner] if we do not hear back from you by the end of today. Please call your HeartStrong Program Advisor, [EA name], at [EA's phone number]."
5. *Escalation level 5 — Phone call to partner*: At the highest escalation level, the EAs first call the patient's partner after non-adherence starts. The EAs mostly made telephone calls to patients' support partners after the seventh day of non-adherence to inform support partners of the adherence problems and enlist their support. In total, the EAs made 46 calls that successfully reached the patients' partners.

Although we did not deliver calls and manual messages regularly like automated messages, it was not uncommon for the patients or the patients' partners to receive them multiple times within a non-adherence sequence. In our main analysis, we consider only the calls that successfully reached the intended recipient. Apart from these five forms of intervention, the EAs also mailed letters to patients in eight cases, and contacted the patient's primary care provider's office by telephone to inform them of adherence problems in three cases. We exclude those interventions from our analysis because they were very rare events and we do not know when the letters were delivered to the patients and if they were received.

As Figure 2 shows, there exists considerable variation in when we delivered the intervention. We did not deliver the intervention in a perfectly consistent manner for several reasons. For automated messages, the inconsistency is primarily because (1) the patients were sometimes unable to receive messages and (2)

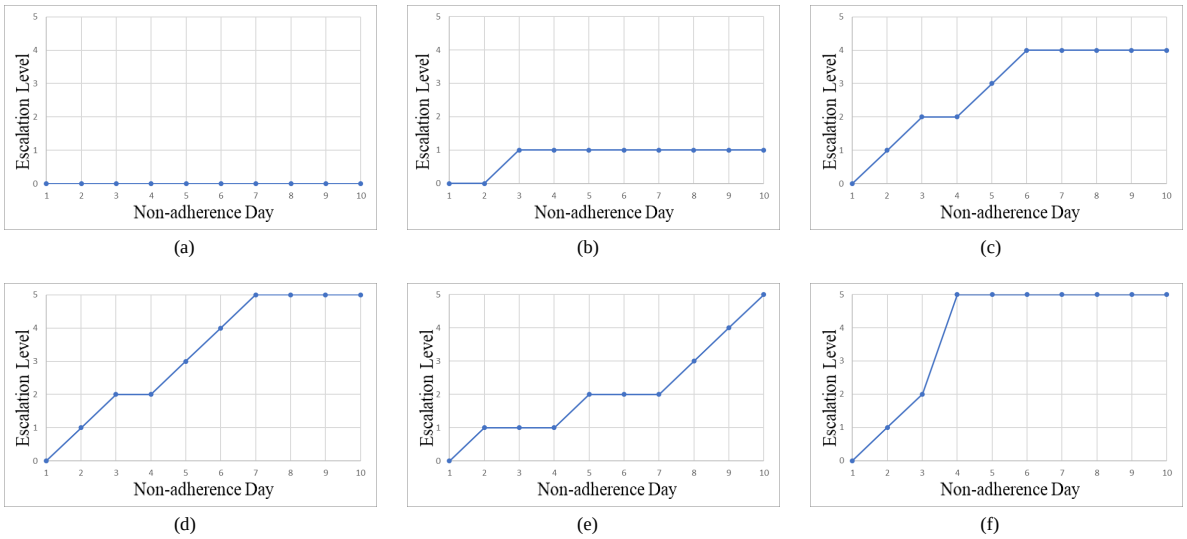
Figure 2 Feedback Distribution over Time



our electronic platform occasionally had errors that prevented it from successfully sending messages. For manual messages and phone calls, we observe considerable variation in delivery timing due to operational constraints such as our care team having limited capacity, patients not being reachable, and the intervention not being delivered on the weekends. Moreover, due to a lack of a strict intervention protocol, we also sometimes delivered manual interventions relatively early within a non-adherence period. These limitations result in varying escalation patterns, which are illustrated in Figure 3. The figure shows six examples of the way the intervention was escalated. In each plot, the marks on the vertical axis indicate current escalation levels, with level 0 meaning no intervention has been delivered. As we can see, there exists significant variation in whether and when each escalation level was reached. Since the variation resulted mostly from operational challenges on our end, we have no reason to believe that the variation is driven by patient-specific factors. Thus, from a research design perspective, the operational challenges created quasi-experiments that we can use to estimate the effect of escalation on adherence.

Since our goal is to estimate the effect of different escalation strategies on the probability of becoming adherent again on a given non-adherence day, we first need to define what characterizes the escalation leading up to that day. We denote the characteristics of escalation that patient i receives on the t^{th} day of non-adherence within the j^{th} non-adherence sequence as $Escalation_{ijt}$. For example, $Escalation_{15,2,5}$ corresponds to the escalation that patient 15 in our sample receives on the fifth day of non-adherence in his or her second non-adherence period. Based on the varying escalation patterns we observe, we describe $Escalation_{ijt}$ using two dimensions:

Figure 3 Examples of Escalation Patterns from Six Different Non-Adherence Sequences



Note. (a) No intervention was delivered, (b) Intervention was escalated to level 1 and stopped being escalated, (c) Intervention was escalated to level 4 and stopped being escalated, (d) Intervention was consistently escalated to level 5, (e) Intervention was escalated to 5, but was slowly escalated at the beginning, (f) Intervention was quickly escalated to level 5, bypassing levels 3 and 4.

1. *Current level of escalation* ($EscLevel_{ijt}$): This captures the intensity of escalation on non-adherence day t. As defined above, levels of escalation range from 0 to 5 with increasing degrees of intensity. The escalation reaches level 1 when the patient first receives an automated message and reaches level 5 when their partner first receives a phone call.
2. *Dynamics of escalation*: This captures the dynamics of escalation leading up to non-adherence day t. The escalation dynamics are further broken up into:
   (a) *Overall escalation dynamics* ($OverallEsc_{ijt}$): $OverallEsc_{ijt}$ is observed from the time non-adherence begins up to non-adherence day t. We quantify $OverallEsc_{ijt}$ as the overall rate of escalation, which is calculated by dividing $EscLevel_{ijt}$ by t. This captures the speed of escalation, which can also be interpreted graphically as the slope of the escalation graph measured from the beginning of the non-adherence sequence to non-adherence day t.
   (b) *Recent escalation dynamics* ($RecentEsc_{ijt}$): Since $OverallEsc_{ijt}$ does not necessarily capture recent escalation dynamics, we use $RecentEsc_{ijt}$ to indicate whether the intervention has been escalated within the past three days. Specifically, we define $RecentEsc_{ijt}$ as follows.

$$RecentEsc_{ijt} = \begin{cases} 1, & \text{if } \begin{cases} EscLevel_{ijt} - EscLevel_{ij,t-2} > 0, & t \geqslant 3 \\ EscLevel_{ijt} > 0, & t < 3 \end{cases} \\ 0, & \text{if } \begin{cases} EscLevel_{ijt} - EscLevel_{ij,t-2} = 0, & t \geqslant 3 \\ EscLevel_{ijt} = 0, & t < 3 \end{cases} \end{cases} \tag{1}$$

In the following section, we incorporate these escalation characteristics into the econometric model that we present and explain how we identify the effect of different escalation strategies on the probability of becoming adherent again.

4.3. Econometric Model

To study the effectiveness of the feedback, we conduct discrete-time survival analysis where we view each non-adherence period as a spell that starts when the patient misses the medications and ends when the patient resumes taking the medications again. Discrete-time survival models are often used to model events that happen in truly discrete time and events that happen in continuous time but are observed in discrete intervals, i.e., interval-censored events. In our case, we observe whether the patient resumed taking the medications daily. Many researchers use continuous-time survival models with discrete-time data and interval-censored data because continuous-time models are easier to implement. In contrast, we use a discrete-time model for two reasons. First, we have many ties in the time when patients became adherent again, especially on the first few days of non-adherence. According to Chalita et al. (2002), one should use a discrete-time model when the proportion of ties is greater than 0.25, which is the case in our data. Second, discrete-time models are less error-prone to implement since our data contain time-varying escalation characteristics that change at high frequency.

To conduct discrete-time survival analysis, we use a complementary log-log (cloglog) model. The cloglog model, which is also known as a discrete-time proportional hazards model, is a mathematically exact time-aggregated version of the continuous-time Cox proportional hazards model (Allison 1982, Prentice and Gloeckler 1978, Jenkins 1995). Like the Cox model, the cloglog model makes no assumption regarding the nature of the hazard function, allowing the baseline probability of becoming adherent on the t^{th} day of non-adherence to take any distributional form. We conduct our survival analysis at the patient-day level where the discrete-time survival function, $S(t|Escalation_{ijt}, X_{ijt})$, is the probability that patient i remains non-adherent for at least t days within non-adherence sequence j; and the discrete-time hazard function, $H(t|Escalation_{ijt}, X_{ijt})$, is the probability that patient i becomes adherent again on non-adherence day t within non-adherence sequence j given that he or she has been non-adherent up until that point. $Escalation_{ijt}$ and X_{ijt} are a vector of escalation characteristics and a vector of control variables, respectively.

Based on the proportional hazards framework, the survival function takes the following form:

$$S(t|Escalation_{ijt}, X_{ijt}) = S_0(t)^{exp(Escalation_{ijt}\beta + X_{ijt}\delta)\nu_i} \tag{2}$$

where $S_0(t)$ the baseline survival function. Since each patient could have multiple non-adherence sequences, we allow observations within each patient to be correlated by introducing a patient-specific frailty, ν_i, which is an unobserved quantity that is log-normally distributed with a mean of one and a variance of σ_ν^2. A frailty,

or a latent random effect, is the same for each patient and is used to describe unobserved heterogeneity among patients (Meyer 1990). Our model is known as a shared-frailty model, which is the survival-analysis analog to random-effects regression models. From Equation 2, we obtain a similar relationship for the complement of the hazard function:

$$1 - H(t|Escalation_{ijt}, X_{ijt}) = (1 - H_0(t))^{exp(Escalation_{ijt}\beta + X_{ijt}\delta)\nu_i} \tag{3}$$

where $H_0(t)$ the baseline hazard on non-adherence day t. We can rewrite Equation 3 in the form of cloglog model as follows (Kalbfleisch and Prentice 2002, Allison 1982).

$$\begin{aligned} cloglog(H(t|Escalation_{ijt}, X_{ijt})) &= ln(-ln(1 - H(t|Escalation_{ijt}, X_{ijt}))) \\ &= \alpha_t + Escalation_{ijt}\beta + X_{ijt}\delta + \mu_i \end{aligned} \tag{4}$$

where $cloglog(\cdot)$ is a complementary log-log link function, and α_t is a constant for non-adherence day t representing the baseline hazard and is equal to $cloglog(H_0(t))$. The log frailty μ_i, or $ln(\nu_i)$, is analogous to random effects in standard regression models and is assumed to be i.i.d., $N(0, \sigma^2_\mu)$. We can determine whether it is necessary to include the frailty by testing the hypothesis that the proportion of total variance contributed by the patient-heterogeneity variance (σ^2_μ) is equal to zero. Similar to the Cox model, we can interpret the exponent of a coefficient as a hazard ratio, which captures a proportional shift in the hazard due to a unit change in the associated covariate given all other factors, including the frailty, being equal.

The main treatment variable in our model is $Escalation_{ijt}$, which consists of $EscLevel_{ijt}$, $OverallEsc_{ijt}$, and $RecentEsc_{ijt}$. We model $EscLevel_{ijt}$ as a factor variable to allow for a non-linear effect. Additionally, we include X_{ijt} to control for seasonality and other factors that potentially influence medication adherence. Specifically, X_{ijt} consists of the following.

1. *Seasonality*: We control for two main seasonality factors: (1) weekends and holidays and (2) days since enrollment. To control for weekends and holidays, we include a binary indicator that equals one if the observation falls on a weekend or a federal holiday. Additionally, since Staats et al. (2017) suggests that individuals' compliance may vary by how long they have been monitored, we control for the number of days the patient had been enrolled in the program.
2. *Financial outcomes*: Our original study used lottery-based incentives to promote medication compliance in addition to providing real-time intervention. Each day, the electronic platform randomly selected a lottery number, which was then compared with the patient's assigned lottery number. If the numbers matched under certain criteria and the patient was adherent the previous day, the patient would receive a monetary prize. Adherent patients were told if they won, and non-adherent patients were told if they would have won had they been adherent, allowing us to leverage regret aversion and anticipated regret. To control for lottery outcomes, we include $Regret_{ijt}$, which is a binary variable that equals 1 if the patient was eligible to win but did not receive the prize due to non-adherence.

3. *Patient characteristics*: Random-effects models allow us to include time-invariant patient characteristics. We include age, gender, PHQ-2 score, baseline Elixhauser score, and whether or not the patient is enrolled in Medicare.
4. *Variation in practices among EAs*: Since the EAs possibly had varying intervention delivery practices that may in turn affect patients' adherence behavior, we include EA fixed effects to control for this.

In our analysis, we aim to separately identify the effect of escalation level and the effect of escalation dynamics. We are able to do so because the correlations among $EscLevel$, $OverallEsc$, and $RecentEsc$ are low. Specifically, we find that the magnitudes of correlation among them are all below 0.5. Moreover, we also aim to disentangle the effect of escalation from the effect of time. Although we control for the baseline hazard rate on non-adherence day t, high correlations between escalation characteristics and non-adherence day may prevent us from reliably identifying the underlying baseline hazard rates and the effect of escalation. To address this, we obtain the correlation coefficients between (1) non-adherence day and each $EscLevel$, (2) non-adherence day and $OverallEsc$, and (3) non-adherence day and $RecentEsc$, and find that the magnitudes of the correlation coefficients are below 0.62, with most being below 0.1.

Even though we do not have a high correlation problem, it is still possible that the effect of escalation level depends on when the escalation level is reached. For example, patients may be less sensitive to escalation level 4 that occurs on non-adherence day 6 than one that occurs on non-adherence day 3 because when patients reach the sixth day of non-adherence, they are naturally more non-adherent and are less likely to respond to intervention. If this were to be true, the fact that higher levels of escalation tend to occur later within a non-adherence sequence could lead to a downward bias in the estimated effect of escalation level. This would also violate the proportionality assumption of the discrete-time hazard model, which requires the effect of escalation level to be identical in every non-adherence day t. To address this, we will test the proportionality hazard assumption by adding the interaction terms between escalation levels and non-adherence day (Singer and Willett 2003, Therneau and Grambsch 2000).

4.4. Results

Table 3 summarizes the relationship between escalation characteristics and the probability of becoming adherent again in terms of hazard ratios. As we are primarily interested in estimating the effect of feedback escalation, we only report the effects of $EscLevel$, $OverallEsc$, and $RecentEsc$. We do not find a statistically significant relationship between the lowest escalation levels that involve only automated messages and the probability of returning to adherence. However, higher escalation levels that involve more customized feedback and a greater degree of personal involvement were associated with greater patient responsiveness. Holding everything else constant, the probability of becoming adherent increased by 28% when the patient started receiving personal phone calls, and increased slightly higher when the patient started receiving manual messages. More remarkably, when the patient's partner started receiving

phone calls, the probability of becoming adherent again increased by more than 35%. Additionally, we find that patients were more likely to become adherent again when the intervention was escalated quickly and when they received a recent escalation. When the patient received an escalation in the past three days, the probability of becoming adherent again increased by 32%. To check for robustness, we also consider voicemails in addition to calls that successfully reached the intended recipient. Although the estimated effects of escalation level 3 (call to patient) and escalation level 5 (call to partner) are slightly smaller after including voicemails, the findings are qualitatively similar.

To further investigate if the escalation levels that only involve automated messages are ineffective in making the patients adherent again or if we simply do not find an effect because of a low signal-to-noise ratio, we perform equivalence tests (Harms and Lakens 2018, Rogers et al. 1993). The first step of the equivalence test is to specify the smallest effect size of interest (SESOI), which is the effect size that we consider too small to be meaningful. To determine what the SESOI should be, we first hypothetically assume that patients did not receive any intervention. Then, we determine by what percentage, on average, the hazard rates would have to increase for the patients to be more likely than not to become adherent when they were currently less likely than not to become adherent. We find the average minimum increase to be 2.882%. Using 2.882% as the SESOI, we perform two one-sided t-tests to determine if we can reject the hypothesis that the effect size is larger than or equal to the SESOI. The test results suggest that we can reject the null hypothesis for both escalation level 1 (p-values < 0.001) and escalation level 2 (p-values < 0.05). Therefore, the effects of escalation levels 1 and 2, if they exist, are likely to be of negligible significance.

Table 3 Estimated Effect of Escalation on the Probability of Becoming Adherent

Variable	Hazard Ratio (SE)
$EscLevel$	
Level 0 – Intervention not yet delivered	(base case)
Level 1 – Automated message to patient	1.01 (0.00)
Level 2 – Automated message to partner	1.01 (0.01)
Level 3 – Phone call to patient	1.28^{***} (0.05)
Level 4 – Manual message to patient	1.29^{**} (0.10)
Level 5 – Phone call to partner	1.36^{**} (0.16)
$OverallEsc$	1.18^{*} (0.09)
$RecentEsc$	1.32^{***} (0.12)

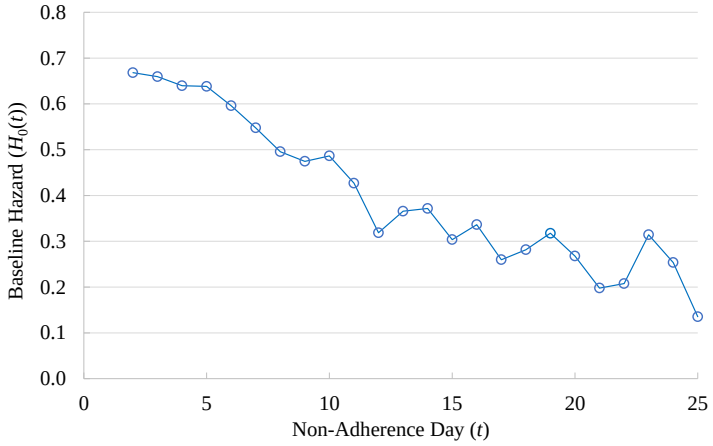
Note. Standard error in parentheses. $^{*}(p < 5\%),^{**}(p < 1\%),^{***}(p < 0.1\%)$.

To determine whether the effect of escalation level depends on non-adherence day, we compare a model that includes interaction terms between escalation levels and non-adherence day to a model that does not. We find that interaction coefficients are not statistically significant. Moreover, by conducting a deviance goodness-of-fit test, we find that the model that allows for a time-dependent effect does not provide a better fit than the proportional hazards model. This suggests that there is insufficient evidence to indicate that

the effect of escalation level depends on non-adherence day, and, therefore, the fact that higher levels of escalation tended to occur later in a non-adherence sequence should not bias the estimates. Furthermore, to determine whether patient-specific frailties are necessary, we conduct a likelihood ratio test to evaluate the null hypothesis that the proportion of the total variance contributed by the patient-heterogeneity variance (σ_μ^2) is zero. The p-value for the likelihood ratio test is less than 0.001. Thus, we can reject the null hypothesis and conclude that patient-specific frailties are important.

Figure 4 shows the baseline hazard for each non-adherence day t, which translates to the probability that the patient becomes adherent again on the t^{th} day of non-adherence given that all covariates are zero. As we can see, the baseline hazard tended to decrease over time, suggesting that patients were less likely to become adherent again the longer they had been non-adherent. We report the full estimation results except for the baseline hazards and patient-specific frailties in Table 11 in the Appendix. In contrast to Staats et al. (2017), which finds that employees were less likely to be compliant the longer they had been monitored, we do not find that the probability of becoming adherent changed the longer the patient had been enrolled in the program. One possible reason for this difference is that, in our study, we not only monitored patients' adherence but also proactively reminded patients to take medications and provided financial incentives to promote adherence throughout the study period. Due to these constant stimulations, patients' likelihood of becoming adherent was likely unaffected by how long they had been in the program. This finding potentially implies that compliance monitoring alone may not be enough to maintain compliance—one also needs to consistently provide patients (or employees) with behavioral feedback and/or incentives to promote compliance.

Figure 4 Baseline Probability of Becoming Adherent Again $(H_0(t))$



The results from our study suggest that, in order to effectively make a previously non-adherent patient adherent again, we should start providing personalized feedback as quickly and as consistently

as possible. Specifically, in our setting, we should start calling the patient immediately after the first day of non-adherence and, if non-adherence continues, escalate to manual messages and calls to the patient's partner successively on the following days. When combining the effect of escalation level and escalation dynamics, we find that receiving a personal phone call on the second day of non-adherence, a manual message on the third day of non-adherence, and a partner call on the fourth day of non-adherence more than doubled the probabilities of becoming adherent during those days.

5. Study B—Impact of Medication Adherence on Readmission

To investigate how behavior change improved health outcomes, we explore the effect of medication adherence on the likelihood of readmission. Specifically, we aim to examine how consistency in medication use in the past affects the readmission probability in the present.

5.1. Data and Statistics

We utilize a panel data set that includes (1) daily data for each patient indicating whether he or she opened each of the pill bottles and (2) daily data of patient readmissions. Consistent with most prior work, we consider as readmissions all-cause inpatient hospitalizations and observation stays, but not emergency room visits. These criteria are also frequently used when hospital performance is assessed. Table 4 provides the distribution of the number of readmissions between the initial discharge and the end of the one-year study period. As the figure shows, approximately 36% of the patients in our cohort were readmitted at least once between their initial hospital discharge and the end of the study period. Apart from medication adherence, patient characteristics may also contribute to the risk of readmission. Table 5 provides summary statistics of patient demographics for all patients in our cohort along with a breakdown by whether or not the patients were readmitted between their initial discharge and the end of the study period.

Table 4 Distribution of the Number of Readmissions between Initial Discharge and the End of Study Period

Number of Readmissions	Percentage
0	64.41%
1	20.00%
2	7.69%
3	3.49%
4	2.15%
5 and above	2.26%

In our study, we analyze the impact of adherence to two medications: statins (which lower cholesterol) and β-blockers (which lower the heart rate). Each pill bottle the patients received was dedicated to either one of the medications. Statins are a class of drugs that help lower cholesterol levels in the blood. β-blockers are usually prescribed to patients with high blood pressure. Statins and β-blockers are generally well-tolerated and nearly universally recommended to patients following heart attack. Although we were able to track

Table 5 Summary Statistics of Patient Demographics by Readmission Profile

	Not Readmitted (N = 628)				Readmitted (N = 347)				All (N = 975)			
Variable	mean	std	min	max	mean	std	min	max	mean	std	min	max
Age	61.03	10.35	23	80	61.68	10.52	28	80	61.26	10.41	23	80
Female	0.31	0.46	0	1	0.40	0.49	0	1	0.34	0.47	0	1
Medicare	0.39	0.49	0	1	0.53	0.50	0	1	0.44	0.50	0	1
PHQ-2 depression score	1.10	1.53	0	6	1.53	1.77	0	6	1.26	1.63	0	6
Baseline Elixhauser score	5.11	9.12	-14	37	8.87	11.08	-13	44	6.46	10.03	-14	44

adherence to each of the two medications separately, we consider adherence to statins and β-blockers combined because the correlation between adherence to the two medications is high. For these medications, we are interested in studying the impact of short- and long-term adherence on readmission. We define short-term adherence as adherence measured over the three days leading up to any given day, and long-term adherence as adherence measured over the 120 days leading up to any given day. We drop the first 120 days of the study for each patient because they do not have long-term adherence information. The adherence measures are binary, taking the value of 1 when the average adherence within a given time window is at least 80% and taking the value of 0 otherwise. The 80% cutoff is widely used in medication adherence literature (Burnier 2019). Given that patients took one dosage of each medication per day, this adherence cutoff implies that they had to take medications for three out of three days to be adherent short-term.

Our main consideration in determining the appropriate lengths of short- and long-term adherence windows was the correlation between short- and long-term adherence. We find that the larger the difference between the lengths of short- and long-term adherence windows is, the lower the correlation between short- and long-term adherence will be. We use three days for short-term adherence and 120 days for long-term adherence because the correlation between 3-day and 120-day adherence is only 0.398. Of course, the correlation would be even lower if we used a longer-length long-term adherence window, e.g., 150 days. However, we would have to drop more observations that do not have long-term adherence information. We believe that, by using three and 120 days, we can balance the need to minimize the correlation with the preservation of information. Our choices of adherence windows are also supported by the LASSO regression, which suggests that 3-day adherence and 120-day adherence are the most contributive adherence measures.

To ensure that dropping the first 120 days of the study does not cause a sample selection issue, we examine whether adherence behavior varied with how long the patients had been in the program. First, we consider the effect of the number of days since enrollment, which is a control variable in Study A, on the probability of becoming adherent. The results indicate that the effect of the number of days since enrollment is not statistically significant (p-value $= 0.409$). Second, we consider the average adherence rates before and after day 120. We find that the average adherence rates before and after day 120 were 0.810 and 0.805, respectively. We also find that the difference is not statistically significant (p-value $= 0.102$). These results

suggest that there exists insufficient evidence to indicate that patients' adherence behavior was different in the first 120 days. Therefore, dropping the first 120 days of the study should not cause a sample selection issue.

Table 6 shows average short- and long-term adherence to statins and β-blockers leading up to days patients were readmitted and average adherence leading up to days patients were not readmitted. We find that average short- and long-term adherence was lower leading up to days patients were readmitted. For both adherence measures, we assume that patients were prescribed a medication from the time of the first fill until the end of the study period. To reduce errors in measuring adherence, we exclude patients whose communication history indicated that they had difficulties setting up their pill bottles or that their pill bottles did not work as intended. We also remove days that patients stayed in the hospital and assume that they were adherent on those days because they were under hospital care.

Table 6 Average Short- and Long-Term Adherence to Statins and β-blockers

Variable	Prior to Day with No Readmission	Prior to Day with Readmission
3-day statin and β-blocker adherence	0.885	0.802
120-day statin and β-blocker adherence	0.880	0.746

5.2. Econometric Model

To study the effect of medication adherence on the risk of readmission, we conduct discrete-time survival analysis where we view each *healthy* period as a spell that starts when the patient is discharged from the hospital and ends when the patient is readmitted. Instead of using the cloglog link function, we use the logit link function, which is another common link function for discrete-time hazard models (Allison 1982, Cox 1972). The logit function allows us to incorporate patient fixed effects, which account for all observable (both available and unavailable in our data) and unobservable patient-specific factors that influence the patient's likelihood of readmission and do not vary over the span of the study period. Incorporating patient fixed effects ensures that our estimates are immune to an omitted variable bias due to unobservable time-invariant factors that are correlated with both medication adherence and readmission. For example, patients who consistently feel unwell may be more likely to adhere to their medications and, because they are likely in poorer health, are also more likely to be readmitted. On the other hand, one can also argue that healthy patients who are less likely to be readmitted tend to have higher motivation and thus are likely to be more adherent. This is known in the literature as the healthy-user effect (Shrank et al. 2011). Patient fixed effects adjust for these potential time-invariant attributes that are also confounding variables.

Although we do not have a high proportion of ties in the failure time and could potentially use a continuous-time model for convenience, we use a discrete-time model because it allows us to extend the model to account for time-varying confounders afterwards. We conduct our analysis at the patient-day

level where the discrete-time hazard function, $H(t|Adherence_{ijt}, Z_{ijt})$, is the probability that patient i is readmitted on the t^{th} day since last discharge within healthy period j. We specify our model as follows:

$$logit(H(t|Adherence_{ijt}, Z_{ijt})) = \alpha_t + Adherence_{ijt}\beta + Z_{ijt}\delta + \gamma_i \tag{5}$$

where $Adherence_{ijt}$ is a vector of short- and long-term adherence measures; α_t is the logit transformation of the baseline hazard ($logit(H_0(t))$); and γ_i is patient i's fixed effect. Z_{ijt} is a vector of control variables, which include (1) a binary variable indicating whether the observation falls on a weekend or a holiday, and (2) the number of previous readmissions, i.e., the number of times the patient had been readmitted prior to the present day. Unlike our original randomized controlled study, this model allows us to evaluate the probability of readmission by leveraging within-patient intertemporal variation in medication adherence.

The fixed-effects logit model is a discrete-time equivalent of the stratified Cox model, i.e., Cox model with each patient treated as a separate stratum (Allison 1996, Allison and Christakis 2006). As opposed to the cloglog and Cox models where the exponent of a coefficient is simply a hazard ratio, the exponent of a coefficient from the logit model is an odds ratio of hazard rates (Cox 1972). However, this converges to a hazard ratio in our case because the probabilities of readmission are generally very small.

To fit the fixed-effects logit model, we use a conditional logistic approach where we group data by patient and calculate the likelihood relative to each patient group, i.e., a conditional likelihood is used (Chamberlain 1980). Through this approach, γ_i's are not directly estimated since the conditional likelihood does not directly involve the patient fixed effects. We estimate the model using robust-cluster standard errors, i.e., standard errors adjusted for clustering of observations within patients, to allow a patient's observations to be correlated. In all our analyses, we check for robustness by varying the length of short-term and long-term adherence windows and by varying the adherence threshold. Specifically, we consider 5 and 7 days as the short- adherence window, and 80 and 100 days as the long-term adherence window. For the adherence threshold, we also consider using 60%, 70%, and 90%.

5.3. Results

Table 7 summarizes the relationship between short- and long-term medication adherence and readmission. The odds ratio for long-term adherence is statistically significant and less than one, suggesting that adhering to statins and β-blockers in the long term was associated with a reduced likelihood of readmission. Specifically, we observe that taking statins and β-blockers at least 80% of the required amount during the past 120 days was associated with a 51% reduction in the odds of readmission. This translates to a similar-sized reduction in the risk of readmission since the probability of readmission is generally very small. Given that patient fixed effects are zero, we find that the average marginal effect of long-term adherence is -0.373%, which is quite significant considering that the probability of readmission is usually lower than 1%.

Table 7 Estimated Effect of Medication Adherence on the Likelihood of Readmission

Adherence Measure	Odds Ratio (SE)
3-day statin and β-blocker adherence	0.75 (0.29)
120-day statin and β-blocker adherence	0.49** (0.12)

Note. Standard error in parentheses. $^{*}(p < 5\%)$,** $(p < 1\%)$,*** $(p < 0.1\%)$.

We report the full estimation results in Table 12 in the Appendix and check for robustness when using different specifications of adherence. As shown in Tables 13 and 14 in the Appendix, the results are robust when we vary the adherence threshold between 60% and 90% and when we vary the short-term adherence window between 3 and 7 days, and the long-term adherence window between 80 to 120 days.

In our analysis, we use patient fixed effects to account for time-invariant unobservable patient attributes. However, there may also exist unobservable time-varying factors that influence both readmission and adherence behavior. For example, patients may be more likely to be adherent at times when they feel unwell or at times when they are healthier and have higher motivation. Although researchers usually view the healthy-user effect as a person-specific time-invariant phenomenon, we do not want to rule out the possibility that it varies with time. To address this, we use an instrumental variable (IV) approach to control for potential endogeneity. As we discuss in detail in Appendix B, we use intervention as IVs for medication adherence and conduct the analysis using a subset of patients who normally had relatively low adherence and needed intervention to stay adherent. Using the IV approach, we find the results to be quite similar to what we obtain using patient fixed effects. Specifically, we find that the effect of long-term adherence is statistically significant while the effect of short-term adherence is not. For long-term adherence, we find the average marginal effect on the probability of readmission to be -0.340%.

6. Readmission Risk Prediction and Its Application in Connected Healthcare

Being able to predict the risk of readmission can help health professionals to effectively intervene with non-adherent patients. We aim to use machine learning to develop a model that predicts readmission risk using recent medication adherence in addition to static risk factors. By utilizing adherence information leading up to each day, one can obtain readmission risk at the patient-day level and use this information when intervening with non-adherent patients. As a care team usually handles a large number of patients and has a limited intervention capacity, it would be helpful for them to be able to identify patients (and patient-days) that are at risk for an imminent readmission and would therefore be prime candidates for the care team's attention.

In this section, we aim to (1) develop a dynamic readmission risk-scoring model that uses medication adherence as predictors (Section 6.1) , (2) compare its performance to a baseline model that excludes medication adherence and only contains baseline risk factors such as patient characteristics (Section 6.2), and (3) apply the dynamic readmission risk-scoring model to our connected health setting to better target non-adherent patients (Section 6.3).

6.1. Model Development

Most traditional readmission prediction models only predict the probability of readmission within a fixed period after discharge using static predictors captured at discharge. Unlike previously developed models, we want to develop a model that predicts patient-day level risk using recent medication adherence in addition to static predictors and compare its performance to a baseline model that excludes medication adherence. In particular, we consider the following models:

1. *Baseline risk-scoring model*: The baseline risk-scoring model only includes the following baseline risk factors:
   (a) *Patient characteristics*: These include age, gender, PHQ-2 score, baseline Elixhauser score, and whether or not the patient is enrolled in Medicare.
   (b) *Seasonality*: Seasonality factors include a categorical variable for month of the year and a binary variable indicating whether the observations falls on a weekend or a holiday.
   (c) *Number of previous readmissions*: This is the number of times the patient had been readmitted prior to the present day.
   (d) *Day since last discharge*: This is the number of days since the patient was last discharged.
2. *Dynamic risk-scoring model*: The dynamic risk-scoring model includes the baseline risk factors presented earlier as well as short- and long-term adherence to statins and β-blockers as defined in Study B.

To train our models, we consider five machine learning methods: (1) logistic regression, (2) decision tree[1], (3) random forest[1], (4) support vector machine[2], and (5) multi-layer perceptron[3]. These classification methods are relatively well-known and are commonly used in the readmission prediction literature. To prepare the data for model training, we standardize all continuous variables to prevent some machine learning algorithms from putting excessive weight on features with large values. Furthermore, we address class imbalance by using synthetic minority over-sampling technique (SMOTE) (Chawla et al. 2002). Since there exist significantly more patient-days without readmission than patient-days with readmission, the machine learning algorithms may bias toward the majority class. To overcome this issue, SMOTE balances the data by creating synthetic observations using k-nearest neighbors. In order to create a synthetic observation, SMOTE finds the k nearest neighbors of each minority observation, selects one of them, and calculates linear interpolations to create a new minority observation in the neighborhood. Research has shown that SMOTE is superior to random oversampling, which is known to increase the likelihood of overfitting.

[1] With Gini impurity criterion.

[2] With radial basis function (RBF) kernel.

[3] With rectified linear unit activation function.

In evaluating the models, we use a five-fold cross validation at the patient level instead of the patient-day level in order to avoid label leaking, i.e., having observations with similar characteristics in both training and validation sets. We use SMOTE within each cross-validation fold after removing the validation sample so that we create synthetic data by interpolating only observations that will not be used for validation.

6.2. Model Evaluation

As a main performance metric, we consider an average area under the ROC curve (AUC) across all five cross-validation folds. Table 8 shows average AUCs along with their standard deviations for the baseline risk models and dynamic risk models that we develop using different machine learning methods. Based on the AUCs, we find that dynamic risk-scoring models outperform baseline risk-scoring models for all classification methods. Furthermore, we find that multi-layer perceptron, which is a class of artificial neural networks (ANN), outperforms non-ANN methods that we consider. The average AUC for the dynamic multi-layer perceptron model is 0.941. This means that, if we take two observations, one with readmission and one without readmission, the model can correctly predict which observation is which 94.1% of the time. Using the dynamic multi-layer perceptron model, we find that the predicted probabilities of readmission on a given day range from 0.01% to 1.98%, with an average of 0.16%. This average is reasonable given that readmission occurred on approximately 0.14% of all patient-days.

Table 8 Average Areas Under the ROC Curve (AUC) for All Models

	Average AUC (SD)	
Method	Baseline Model	Dynamic Model
Logistic regression	0.740 (0.015)	0.884 (0.036)
Decision tree	0.709 (0.051)	0.712 (0.073)
Random forest	0.817 (0.020)	0.909 (0.022)
Support vector machine	0.767 (0.074)	0.856 (0.082)
Multi-layer perceptron	0.843 (0.017)	0.941 (0.021)

To further compare the baseline risk model and the dynamic risk model, we focus on the models developed using multi-layer perceptron and consider a hypothetical scenario where the care team has a limited capacity to intervene with patients to prevent hospital readmissions, e.g., sending medical professionals to patients' homes. We suppose that the care team can target only $c\%$ of all patient-days using one of the three approaches: (1) using no information from a predictive model, (2) using information from the baseline risk model, and (3) using information from the dynamic risk model. If using the first approach, the care team will randomly target $c\%$ of patients each day. If using the second or third approach, the care team will target the patient when his or her predicted risk is higher than $(100-c)^{\text{th}}$ percentile.

To evaluate the three approaches, we define *patient-day at risk* as patient-day with readmission or prior to readmission, which accounts for 0.28% of all patient-days, and consider the following performance measures:

$$Yield = \frac{\text{Number of patient-days at risk we target}}{\text{Number of all patient-days at risk}} \tag{6}$$

$$Precision = \frac{\text{Number of patient-days at risk we target}}{\text{Number of all patient-days we target}} \tag{7}$$

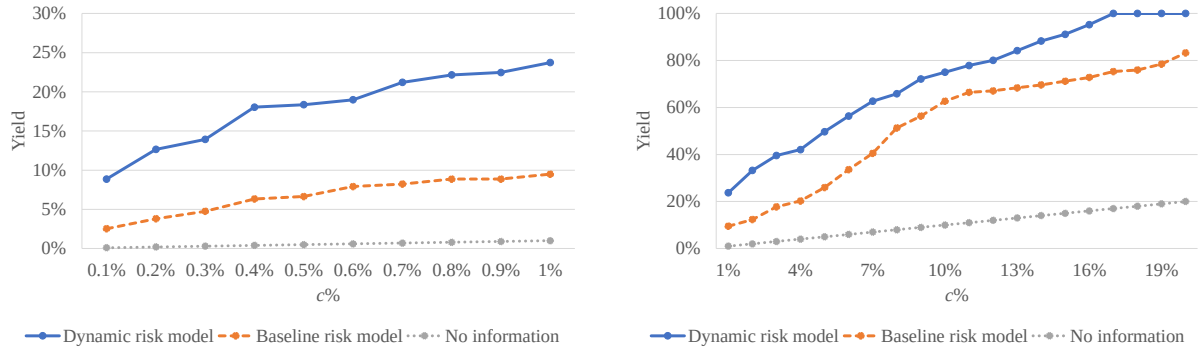
We calculate yield and precision using the validation sample across all cross-validation folds. As seen in Figures 5 and 6, we obtain higher yield and precision when we use the dynamic risk model as opposed to the baseline model, and when we use the baseline risk model as opposed to not using a predictive model at all. Figure 5 shows that yield increases as capacity increases. We find that, if we have the capacity to target 17% of patient-days using the dynamic model, we will be able to successfully target all patient-days at risk. In contrast, if we use the baseline model or use no information from a predictive model, we will be able to successfully target only 75.32% and 17% of the patient-days at risk, respectively. Figure 6 shows that precision generally decreases as we target more patient-days. If we have the capacity to target only 0.1% of patient-days using the dynamic model, as many as 24.56% of the patient-days we target will be patient-days at risk. In contrast, if we use the baseline model or use no information from a predictive model, only 7.02% and 0.28% of the patient-days we target will be patient-days at risk, respectively. We also vary the definition of patient-day at risk and show the results in Figures 9 and 10 in the Appendix. As expected, although yield and precision vary according to the definition of patient-day at risk, we always obtain highest yield and precision when using the dynamic risk model.

Our finding suggests that the variation in medication adherence over time provides useful information about the readmission risk and enhances the quality of prediction in addition to baseline risk factors. We note that the clinical value of adherence-based predictive models depends critically on the real time availability of adherence data.

6.3. Application of Dynamic Readmission Risk-Scoring Model in Connected Healthcare

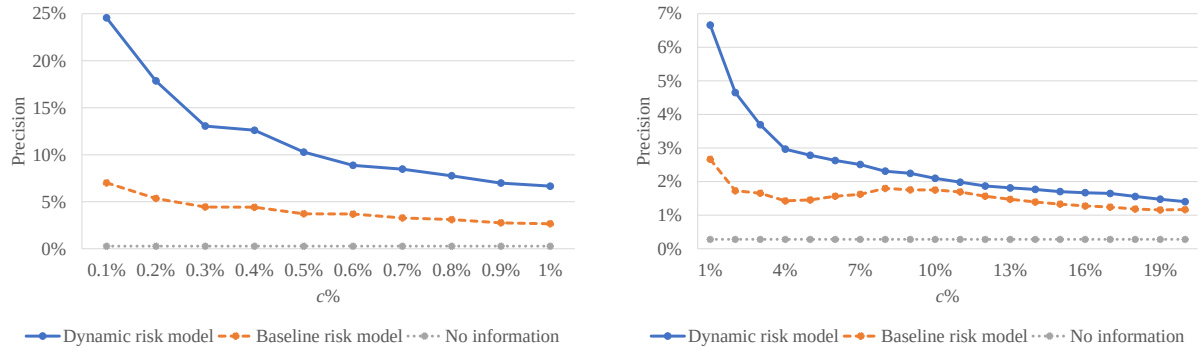
Predictive analytics can benefit connected healthcare by helping to identify patients who really need to change their behavior to avoid adverse health outcomes. In our setting, the dynamic readmission risk-scoring model can help identify patients who are at the highest risk of readmission and have the greatest need to be adherent. We investigate the benefit of using the dynamic risk-scoring model in our setting through counterfactual simulations, which link intervention to readmission by using medication adherence as a mediator. We focus on the multi-layer perceptron model and numerically examine the impact of different intervention strategies, which do and do not consider predicted readmission risks when delivering adherence intervention. The strategies that we consider are as follows.

1. *Strategy A—Adherence-maximizing strategy*: The adherence-maximizing strategy aims to quickly make patients adherent again once they become non-adherent. We define the adherence-maximizing strategy based on the findings from Study A. The strategy involves calling the patient on the second day of non-adherence, sending a manual message on the third day of non-adherence, and calling the patient's partner on the fourth day of non-adherence.



(a) $0.1 \geqslant c \geqslant 1$ (b) $1 \geqslant c \geqslant 20$

Figure 5 Yield Obtained Using Different Readmission Intervention Approaches



(a) $0.1 \geqslant c \geqslant 1$ (b) $1 \geqslant c \geqslant 20$

Figure 6 Precision Obtained Using Different Readmission Intervention Approaches

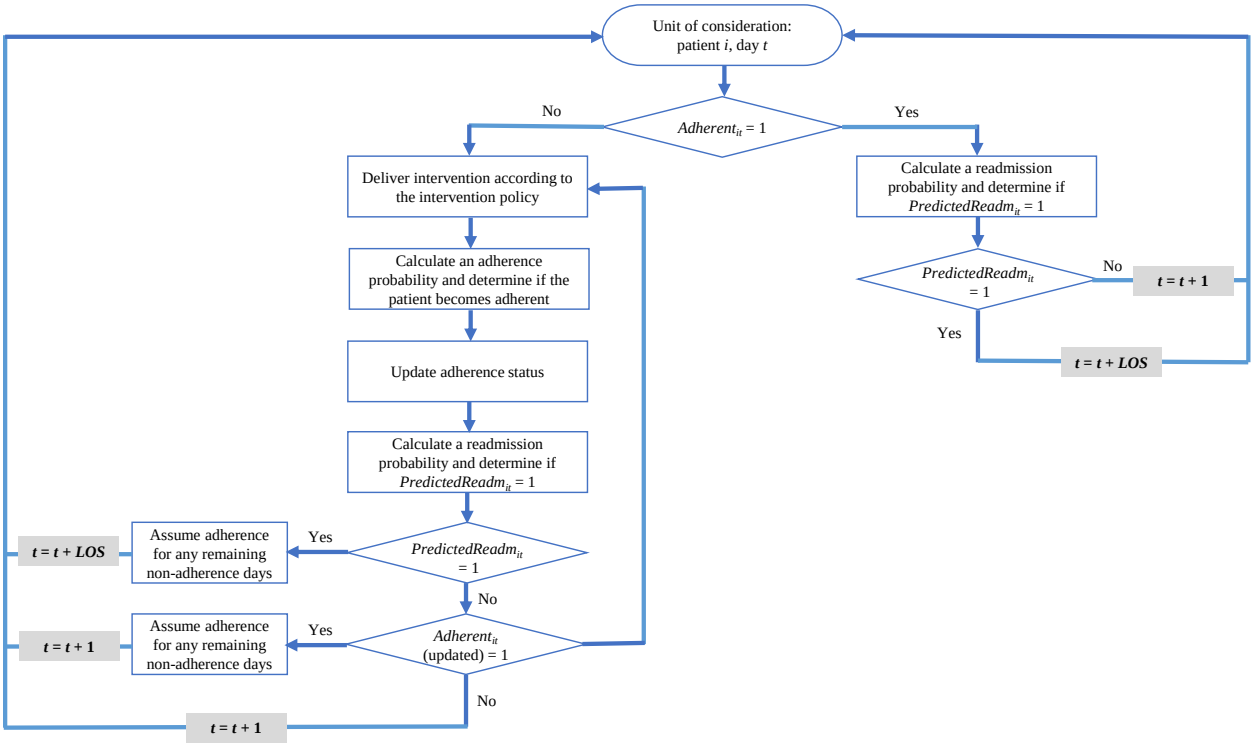
2. *Strategy B—Adherence-maximizing strategy with readmission risk prioritization*: We use the adherence-maximizing strategy, but the strategy is only triggered when the predicted readmission risk is greater than or equal to the n^{th} percentile. We consider four values of n in our main analysis: 80, 85, 90, and 95.
3. *Strategy C—No intervention*: We do not deliver any intervention.

For each intervention strategy, we are interested in four outcome metrics: (1) total number of calls and manual messages delivered, (2) average individual adherence rate, (3) number of patients who are readmitted at least once, and (4) total number of readmissions. We use the empirical numbers that we observe in our experiment as a baseline.

Our counterfactual simulations link intervention to medication adherence using the estimated proportional hazards model from Study A, and link medication adherence to readmission using the dynamic risk-scoring model. The unit of consideration in our simulation is patient i-day t. We consider 930 patients in our focus cohort. For each patient, t starts from 120 until the patient's last day in the study. We illustrate

a simulation flow for patient i in Figure 7 and present our simulation parameters in Table 15 in the Appendix. In Figure 7, $Adherent_{it}$ is a binary variable that equals one when patient i takes all of his or her medications on day t and, at the beginning of the simulation, takes the actual observed value in the data. When the patient is non-adherent, we calculate the probability that the patient will become adherent again using the proportional hazards model from Study A, which takes into account the intervention the patient receives. Based on that probability, we determine whether the patient becomes adherent again using random sampling. $PredictedReadm_{it}$ is a binary variable that equals one when patient i is readmitted on day t. We calculate the probability of readmission using the dynamic risk-scoring model and determine whether the patient is readmitted using random sampling. If the patient is readmitted, he or she will be in the hospital for LOS days, where LOS follows the empirical distribution shown in Table 15. We replicate the simulation 1,000 times and obtain the averages for the four outcomes of interest.

Figure 7 Simulation Flow



Note. (1) This chart depicts a simulation flow for patient *i*. (2) *t* starts from 120 until patient *i*'s last day in the study.

We report the simulation results for different intervention strategies in Table 9 and, to visualize the results, plot the total number of readmissions against the number of calls and manual messages delivered in Figure 8. We find that although Strategy A, which is the most aggressive strategy, can increase adherence to as high as 98.36%, it may not be a practical way to reduce readmissions. While it is true that high medication adherence leads to lower readmission risk, it may be unnecessary to increase medication adherence among

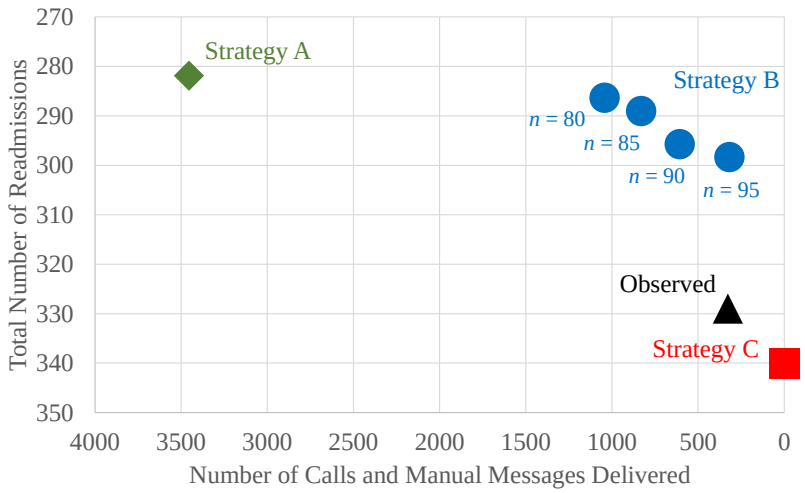
patients who already have relatively low readmission risk and are unlikely to be readmitted. As Figure 8 shows, compared to Strategy A, Strategy B leads to a significant decrease in the total number of calls and manual messages but only a small increase in the total number of readmissions. When n equals to 80 (i.e., we intervene with patients only when the predicted risk is greater than or equal to the 80th percentile), Strategy B leads to a 70% decrease in the total number of calls and manual messages but only a 2% increase in the total number of readmissions compared to Strategy A. Furthermore, we find that, compared to the actual case we observe, Strategy B with n equal to 95 yields a comparable number of calls and manual messages but a significantly lower number of readmissions. Specifically, when using the strategy, we deliver 3% fewer calls and manual messages and obtain 10% fewer readmissions compared to what we empirically observe in the experiment. These results suggest that, by utilizing predicted risk information, the care team can effectively allocate its capacity to the patient-days that really need intervention.

Table 9 Simulation Results for Different Intervention Strategies

	Scenario				
Outcome	Observed	Strategy A	Strategy B ($n = 80$)	Strategy B ($n = 95$)	Strategy C
Total number of manual interventions	327	3453.9	1043.2	318.4	0
Call to patient	210	3190.5	960.1	294.0	0
Manual message to patient	83	239.4	74.8	21.1	0
Call to partner	34	24.1	8.2	3.2	0
Average individual adherence rate	85.50%	98.36%	92.03%	86.94%	81.43%
Total number of readmissions	329	281.9	286.3	297.3	340.0
Number of patients readmitted at least once	288	259.8	261.2	266.2	290.6

Note. Results for Strategy B, $n = 85$ and 90, are show in Table 16 in the Appendix.

Figure 8 Numbers of Manual Interventions and Readmissions for Different Intervention Strategies



Our simulation findings not only highlight the importance of predictive analytics in connected healthcare but also provide a possible explanation for why we did not observe a significant reduction in readmissions in our original randomized controlled trial. Since we could deliver only a limited number of interventions and did not consider readmission risks, it is possible that we delivered many interventions to low-risk patient-days and overlooked high-risk patient-days. As Figure 8 shows, we find that the observed number of readmissions is very close to the number of readmissions we obtain in a simulated scenario where no intervention is delivered. To further investigate this, we apply the readmission risk-scoring model to the observed data and find that more than half of the patient-days that received manual messages or calls had readmission risks below the 59^{th} percentile. Moreover, only 11.42% of the patient-days that received manual messages or calls had readmission risks above the 95^{th} percentile. This evidence supports the explanation that we may have not delivered enough interventions to patient-days with high readmission risk and that our intervention would have been more effective if it had been based on dynamic risk scoring.

7. Conclusions and Managerial Insights

This paper studies the effectiveness of a connected health system that aimed to reduce the number of readmissions through better medication adherence and examines the benefit of predictive analytics in connected healthcare. By utilizing micro-level intervention and adherence data, we find that patients are significantly more likely to become adherent again when they receive high levels of intervention that involve personalized feedback and when the intervention is escalated quickly and continuously. Our findings highlight the importance of personal involvement and speed when delivering connected health intervention, which has not been widely explored by healthcare researchers. Although the marketing and psychology literature has documented the benefit of personal involvement in changing customers' behavior (e.g., Gordon et al. 1998), not much has been studied in the context of healthcare.

To investigate the extent to which behavior change could improve patients' health outcomes, we explore the effect of short- and long-term medication adherence to two crucial medications on readmission risk. We find that, for patients with cardiovascular disease, long-term adherence to statins and β-blockers is associated with a 51% reduction in the odds of being readmitted on any given day. Although we do not find a significant effect of short-term adherence, the lack of significance does not suggest that short-term medication adherence is not important. Rather, our finding implies that, in order to significantly benefit from adherence, patients need to consistently maintain their adherence over a long period. The benefit of long-term adherence is supported by the view in the medical and behavioral compliance literature, which suggests that long-term adherence to medications and treatments is crucial to maintaining good health (Sabate 2003). Most importantly, our findings emphasize the importance of connected healthcare as a tool to improve compliance in the long term by continuously connecting with patients over a long period, as opposed to traditional care where such connectivity is not possible.

In addition to studying the impact of the intervention on medication adherence and the impact of adherence on readmission, we develop a dynamic readmission-risk scoring model that considers real-time medication adherence and find that the model outperforms a baseline risk model that does not consider medication adherence. The results suggest that real-time behavior information can be useful in predicting health outcomes. Although real-time behavior data were not extensively available in the past, they are becoming more widely available thanks to an increasing availability of connected health devices.

Lastly, we examine the benefit of using the dynamic risk model in connected healthcare when the care team has a limited capacity. Using counterfactual simulations, we find that, when using an intervention strategy that prioritizes patient-days with highest readmission risk, we can achieve a significantly lower number of readmissions than we would obtain without considering readmission risk while maintaining the same amount of effort. The dynamic risk model helps the care team allocate their resources to high-risk patients who are non-adherent and have the greatest need to stay adherent. Our findings underscore the value of predictive analytics in connected healthcare. Since connected health technologies allow care providers to conveniently connect with patients, it also enables them to oversee more patients than they traditionally could. With predictive analytics, care providers can optimally allocate their limited resources to achieve the desired outcome. Predictive analytics is not only useful for connected healthcare, but also for other industries that want to implement a connected strategy and make the most out of their limited resources. For example, in financial advising, instead of contacting consumer investors and asking them to rebalance their account every time their asset allocation changes or their cash balance is low, investment firms can use predictive models to target only the investors whose portfolio health is more likely to be affected if they do not rebalance their account.

There exist a number of opportunities for future work. To better understand the effectiveness of connected healthcare, future research could explore whether and how different patient demographics respond differently to intervention. Since it is possible that some patients are more sensitive to certain types of intervention than others, understanding these differences can help health professionals design an effective intervention strategy to target different demographics. We hope that by providing the data that we obtained as part of a connected healthcare experiment to the research community, future research can take advantage of the clinical interventions we performed. Beyond using the data that we collected, we also believe that future research should conduct a randomized controlled trial in which adherence behavior is tracked in both the control and intervention groups, but feedback is delivered only to intervention patients. This would improve our understanding on how exactly patients react to feedback and reminders. Moreover, future connected healthcare studies could conduct a randomized controlled trial that utilizes a predictive model similar to what we propose to better understand how predictive analytics can actually benefit connected healthcare.

Acknowledgments

The randomized controlled trial was funded with a grant from Center for Medicare & Medicaid Innovation (CMS), Health Care Innovation Award 1C1CMS331009.

References

Agency for Healthcare Research and Quality (2014) Statistical brief #172. URL https://www.hcup-us.ahrq.gov/reports/statbriefs/sb172-Conditions-Readmissions-Payer.pdf, accessed June 27, 2019.

Allison PD (1982) Discrete-time methods for the analysis of event histories. *Sociol. Methodol.* 13:61–98.

Allison PD (1996) Fixed-effects partial likelihood for repeated events. *Sociol. Method. Res.* 25(2):207–222.

Allison PD, Christakis N (2006) Fixed effects methods for the analysis of non-repeated events. *Sociol. Methodol.* 36:155–172.

Angrist JD, Pischke JS (2013) *Mostly harmless econometrics: an empiricists companion* (Cram101 Publishing).

Angst CM, Devaraj S, Queenan CC, Greenwood B (2011) Performance effects related to the sequence of integration of healthcare technologies. *Production Oper. Management* 20(3):319–333.

Asch DA, Muller RW, Volpp KG (2012) Automated hovering in health care—watching over the 5000 hours. *N. Engl. J. Med.* 367:1–3.

Bavafa H, Hitt LM, Terwiesch C (2018) The impact of e-visits on visit frequencies and patient health: Evidence from primary care. *Management Sci.* 64(12):5461–5480.

Berry Jaeker JA, Tucker AL (2017) Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Sci.* 63(4):1042–1062.

Bosworth HB, Granger BB, Mendys P, Brindis R, Burkholder R, Czajkowski SM, Daniel JG, Ekman I, Ho M, Johnson M, Kimmel SE, Liu LZ, Musaus J, Shrank WH, Buono EW, Weiss K, Granger CB (2011) Medication adherence: A call for action. *Amer. Heart J.* 162(3):412–424.

Bradley EH, Curry L, Horwitz LI, Sipsma H, Wang Y, Walsh MN, Goldmann D, White N, Piña IL, Krumholz HM (2013) Hospital strategies associated with 30-day readmission rates for patients with heart failure. *Circ. Cardiovasc. Qual. Outcomes* 6(4):444–450.

Burnier M (2019) Is there a threshold for medication adherence? lessons learnt from electronic monitoring of drug adherence. *Front. Pharmacol.* 9:1540.

Cameron AC, Trivedi PK (1998) *Regression analysis of count data* (Cambridge University Press, Cambridge, England).

Campbell D, Frei F (2010) Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Sci.* 56(1):4–24.

Caulfield BM, Donnelly SC (2013) What is Connected Health and why will it change your practice? *QJM-Internat. J. Med.* 106(8):703–707.

Chalita LVAS, Colosimo EA, Demetrio CGB (2002) Likelihood approximations and discrete models for tied survival data. *Commun. Stat. Theory Methods* 31(7):1215–1229.

Chamberlain G (1980) Analysis of covariance with qualitative data. *Rev. Econ. Stud.* 47(1):225–238.

Chan CW, Green LV, Lekwijit S, Lu L, Escobar G (2019) Assessing the impact of service level when customer needs are uncertain: An empirical investigation of hospital step-down units. *Management Sci.* 65(2):751–775.

Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, Phillips CO, Hodshon BV, Cooper LS, Krumholz HM (2010) Telemonitoring in patients with heart failure. *N. Engl. J. Med.* 363(24):2301–2309.

Chaudhry SI, Phillips CO, Stewart SS, Riegel B, Mattera JA, Jerant AF, Krumholz HM (2007) Telemonitoring for patients with chronic heart failure: A systematic review. *J. Cardiac Failure* 13(1):26–62.

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* 16:321–357.

Chouvarda IG, Goulis DG, Lambrinoudaki I, Maglaveras N (2015) Connected health and integrated care: Toward new models for chronic disease management. *Maturitas* 82(1):22–27.

Cleland JGF, Louis AA, Rigby AS, Janssens U, Balk A (2005) Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death: The trans-european network-home-care management system (TEN-HMS) study. *Journal of the American College of Cardiology* 45(10):1654 – 1664.

Cox DR (1972) Regression models and life-tables. *J. R. Stat. Soc. Series B Stat. Methodol.* 34(2):187–220.

Devaraj S, Ow TT, Kohli R (2013) Examining the impact of information technology and patient flow on healthcare performance: A theory of swift and even flow (TSEF) perspective. *J. Oper. Management* 31(4):181–192.

DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW (2002) Patient adherence and medical treatment outcomes: A meta-analysis. *Med. Care* 40(9):794–811.

Gordon ME, McKeage KK, Fox MA (1998) Relationship marketing effectiveness: The role of involvement. *Psychol. Mark.* 15(5):443–459.

Greene W (2004) The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econom. J.* 7(1):98–119.

Greene WH (2012) *Econometric Analysis* (Prentice Hall, Upper Saddle River, NJ), 7th edition.

Gurvich I, O'Leary K, Wang L, Mieghem JAV (2019) Collaboration, interruptions and changeover times: Workflow model and empirical study of hospitalist charting. *Manufacturing and Service Oper. Management* Forthcoming.

Han E, Suh D, Lee SM, Jang S (2014) The impact of medication adherence on health outcomes for chronic metabolic diseases: A retrospective cohort study. *Res. Soc. Admin. Pharm.* 10:87–98.

Harms C, Lakens D (2018) Making 'null effects' informative: statistical techniques and inferential frameworks. *J. Clin. Transl. Res.* 3(Suppl 2):382–393.

Ho PM, Bryson CL, Rumsfeld JS (2009) Medication adherence. *Circulation* 119(23):3028–3035.

Ibanez MR, Clark JR, Huckman RS, Staats BR (2018) Discretionary task ordering: Queue management in radiological services. *Management Sci.* 64(9):4389–4407.

Jenkins SP (1995) Easy estimation methods for discrete-time duration models. *Oxf. Bull. Econ. Stat.* 57:129–136.

Jonasson JO, Boutilier J, Yoeli E (2020) Digital differentiated adherence support: Improving tuberculosis treatment adherence in Kenya. *Working Paper, MIT Sloan School of Management, Cambridge, MA* .

Jonasson JO, Deo S, Gallien J (2017) Improving HIV early infant diagnosis supply chains in Sub-Saharan Africa: Models and application to Mozambique. *Oper. Res.* 65(6):1479–1493.

Kalbfleisch JD, Prentice RL (2002) *The Statistical Analysis of Failure Time Data, Second Edition* (John Wiley and Sons, New York).

Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S (2013) Risk prediction models for hospital readmission: A systematic review. *JAMA* 306(15):1688–1698.

Kim SH, Chan CW, Olivares M, Escobar G (2015) ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Sci.* 61(1):19–38.

Kroenke K, Spitzer RL, Williams JB (2003) The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Med. Care* 41(11):1284–1292.

Kvedar J, Coye MJ, Everett W (2014) Connected health: A review of technologies and strategies to improve patient care with telemedicine and telehealth. *Health Affairs* 33(2):194–199.

Lam WY, Fresco P (2015) Medication adherence measures: an overview. *Biomed. Res. Int.* 217047.

Liu N, Finkelstein SR, Kruk ME, Rosenthal D (2018) When waiting to see a doctor is less irritating: Understanding patient preferences and choice behavior in appointment scheduling. *Management Sci.* 64(5):1975–1996.

Maeng DD, Starr AE, Tomcavage JF, Sciandra J, Salek D, Griffith D (2014) Can telemonitoring reduce hospitalization and cost of care? a health plan's experience in managing patients with heart failure. *Population Health Management* 17(6):340–344.

McDermott MM, Schmitt B, Wallner E (1997) Impact of medication nonadherence on coronary heart disease outcomes. *Arch. Intern. Med.* 157:1921–1929.

McManus D, Pipkin SS, Whooley MA (2005) Screening for depression in patients with coronary heart disease (data from the heart and soul study). *Amer. J. Cardiol.* 96(8):1076–1081.

Meyer BD (1990) Unemployment insurance and unemployment spells. *Econometrica* 58(4):757–782.

Min X, Yu B, Wang F (2019) Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: A case study on COPD. *Sci. Rep.* 9(1):2362.

Moon K, Bimpiki K, Mendelson H (2018) Randomized markdowns and online monitoring. *Management Sci.* 64(3):1271–1290.

Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, Agoritsas T, Mistry N, Iorio A, Jack S, Sivaramalingam B, Iserman E, Mustafa RA, Jedraszewski D, Cotoi C, Haynes RB (2014) Interventions for enhancing medication adherence. *Cochrane Database Syst. Rev.* (11).

Pierce L, Snow DC, McAfee A (2015) Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Sci.* 61(10):2299–2319.

Powell A, Savin S, Savva N (2012) Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing and Service Oper. Management* 14(4):512–528.

Prentice RL, Gloeckler LA (1978) Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 34(1):57–67.

Raghupathi W, Raghupathi V (2014) Big data analytics in healthcare: promise and potential. *Health Inf. Sci. Syst.* 2(3).

Rajan B, Tezcan T, Seidmann A (2019) Service systems with heterogeneous customers: Investigating the effect of telemedicine on chronic care. *Management Sci.* 65(3):1236–1267.

Rogers JL, Howard KI, Vessey JT (1993) Unemployment insurance and unemployment spells. *Psychol. Bull.* 113(3):553–565.

Sabate E (2003) *Adherence to long-term therapies: evidence for action* (World Health Organization, Geneva, Switzerland).

Shrank WH, Patrick AR, Brookhart MA (2011) Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. *J. Gen. Intern. Med.* 26(5):546–550.

Shulan M, Gao K, Moore CD (2013) Predicting 30-day all-cause hospital readmissions. *Health Care Management Sci.* 16(2):167–175.

Siggelkow N, Terwiesch C (2019) *Connected Strategy: Building Continuous Customer Relationships for Competitive Advantage* (Harvard Business Press, Cambridge, MA).

Singer JD, Willett JB (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Oxford University Press, New York).

Staats BR, Dai H, Hofmann D, Milkman KL (2017) Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Sci.* 63(5):1271–1656.

Tan TF, Netessine S (2019) At your service on the table: Impact of tabletop technology on restaurant operations. *Management Sci.* Forthcoming.

Therneau T, Grambsch P (2000) *Modelling Survival Data: Extending the Cox Model* (Springer, New York).

Trappenburg JCA, Niesink A, de Weert-van Oene GH, van der Zeijden H, van Snippenburg R, Peters A, Lammers JJ, Schrijvers AJP (2008) Effects of telemonitoring in patients with chronic obstructive pulmonary disease. *Telemed. e-Health* 14(2):138–146.

van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ (2009) A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med. Care* 47(6):626–633.

Volpp KG, Troxel AB, Mehta SJ, Norton L, Zhu J, Lim R, Wang W, Marcus N, Terwiesch C, Caldarella K, Levin T, Relish M, Negin N, Smith-McLallen A, Snyder R, Spettell CM, Drachman B, Kolansky D, Asch DA (2017) Effect of electronic reminders, financial incentives, and social support on outcomes after myocardial infarction: The HeartStrong randomized clinical trial. *JAMA Intern. Med.* 177(8):1093–1101.

Wang G, Li J, Hopp WJ, Fazzalari FL, Bolling SF (2019) Using patient-specific quality information to unlock hidden healthcare capabilities. *Manufacturing and Service Oper. Management.* 21(3):479–711.

Watson AJ, Kvedar JC, Rahman B, Pelletier AC, Salber G, Grant RW (2009) Diabetes connected health: a pilot study of a patient- and provider-shared glucose monitoring web application. *J. Diabetes Sci. Technology* 3(2):345–352.

Yu AP, Yu YF, Nichol MB (2010) Estimating the effect of medication adherence on health outcomes among patients with type 2 diabetes—an application of marginal structural models. *Value Health* 13(8):1038–1045.

Appendix A: Supplementary Tables and Figures

Table 10 Data Description

Data	Description
Adherence information	Daily information on whether the patient opened each of the pill bottles.
Intervention information	Daily information on whether our care team delivered each type of intervention to the patient, the patient's partner, or the patient's provider's office.
Financial incentive outcomes	Daily lottery results, indicating whether the patient was eligible to win but did not receive the prize due to non-adherence.
Readmission	Readmissions include all-cause inpatient hospitalizations and observation stays, but not emergency room visits. It is a dummy variable, coded 1 if readmission criteria are met and 0 otherwise.
Age	Patient's age. It is a continuous variable, coded as piecewise linear spline variables with knots at its 50^{th} and 80^{th} percentiles to account for potential nonlinear effects.
Gender	Patient's gender. It is a dummy variable, coded as 1 if the patient is female and 0 otherwise.
Medicare enrollment	Information on whether the patient is enrolled in Medicare, a national health insurance program in the United States. It is a dummy variable, coded as 1 if the patient is enrolled and 0 otherwise.
PHQ-2 depression score	Patient's depression score from Patient Health Questionnaire-2 (PHQ-2), which inquires about the frequency of depressed mood and anhedonia over the past two weeks. The score ranges from 0 to 6 with higher scores translating to higher chances of having depression. We asked patients to complete PHQ-2 once at the time of enrollment.
Baseline Elixhauser score	Patient's Elixhauser comorbidity index, which measures comorbidities of patients based on the International Classification of Diseases (ICD) diagnosis code. We assigned the score to patients once at the beginning of the study period using pre-enrollment data. The score is a continuous variable, coded as piecewise linear spline variables with knots at its 50^{th} and 80^{th} percentiles to account for potential nonlinear effects.
Days since enrollment	The number of days since enrollment.
Weekend and holiday	Indicator for weekend and holiday. It is a dummy variable, coded as 1 if the day falls on a weekend or a federal holiday.

Table 11 Full Estimation Results for Study A

Variable	Hazard Ratio (SE)
$EscLevel$	
Level 0 – Intervention not yet delivered	(base case)
Level 1 – Automated message to patient	1.01 (0.00)
Level 2 – Automated message to partner	1.01 (0.01)
Level 3 – Phone call to patient	1.28^{***} (0.05)
Level 4 – Manual message to patient	1.29^{**} (0.10)
Level 5 – Phone call to partner	1.36^{**} (0.16)
$OverallEsc$	1.18^{*} (0.09)
$RecentEsc$	1.32^{***} (0.12)
$Regret$	1.04 (0.03)
Days since enrollment	1.00 (0.06)
Weekend and holiday	0.91 (0.34)
EA 1	(base case)
EA 2	1.05 (0.07)
EA 3	0.87 (0.32)
EA 4	0.94 (0.35)
Age [<62]	1.00 (0.00)
Age [62-71]	1.02 (0.01)
Age [>71]	0.98^{**} (0.00)
Female	1.04^{*} (0.01)
Baseline Elixhauser score [<5]	0.87 (0.33)
Baseline Elixhauser score [5-15]	0.92 (0.34)
Baseline Elixhauser score [>15]	0.99^{*} (0.00)
PHQ-2 score	0.41^{*} (0.19)
Medicare	0.96 (0.04)
Observations	40,250
Wald χ^2_{185}	1319.14 ($p<0.1\%$)

Note. Standard error in parentheses. $^{*}(p<5\%)$,$^{**}(p<1\%)$,$^{***}(p<0.1\%)$. α_t's are not reported.

Table 12 Full Estimation Results for Study B

Variable	Odds Ratio (SE)
3-day statin and β-blocker adherence	0.75 (0.29)
120-day statin and β-blocker adherence	0.49^{**} (0.12)
Number of previous readmissions	3.02^{***} (0.61)
Weekend and holiday	1.17 (0.93)
Number of observations	67,872
Pseudo R^2	0.283

Note. Standard error in parentheses. $^{*}(p<5\%)$,$^{**}(p<1\%)$,$^{***}(p<0.1\%)$. α_t's are not reported.

Table 13 Robustness Test for Study B: Varying Adherence Threshold

Adherence threshold	Odds Ratio (SE)
	Short-term statin and β-blocker adherence
60%	0.81 (0.44)
70%	0.75 (0.34)
80%	0.75 (0.29)
90%	0.73 (0.29)
	Long-term statin and β-blocker adherence
60%	0.54* (0.20)
70%	0.48** (0.13)
80%	0.49** (0.12)
90%	0.46** (0.13)

Note. Standard error in parentheses. $^{*}(p < 5\%)$,$^{**}(p < 1\%)$,$^{***}(p < 0.1\%)$.
We vary the threshold for short- and long-term adherence simultaneously.

Table 14 Robustness Test for Study B: Varying Short- and Long-Term Adherence Windows

Adherence window	Odds Ratio (SE)
	Short-term statin and β-blocker adherence
3-day adherence	0.75 (0.29)
5-day adherence	0.72 (0.32)
7-day adherence	0.80 (0.36)
	Long-term statin and β-blocker adherence
80-day adherence	0.53* (0.23)
100-day adherence	0.50** (0.15)
120-day adherence	0.49** (0.12)

Note. Standard error in parentheses. $^{*}(p < 5\%)$,$^{**}(p < 1\%)$,$^{***}(p < 0.1\%)$.

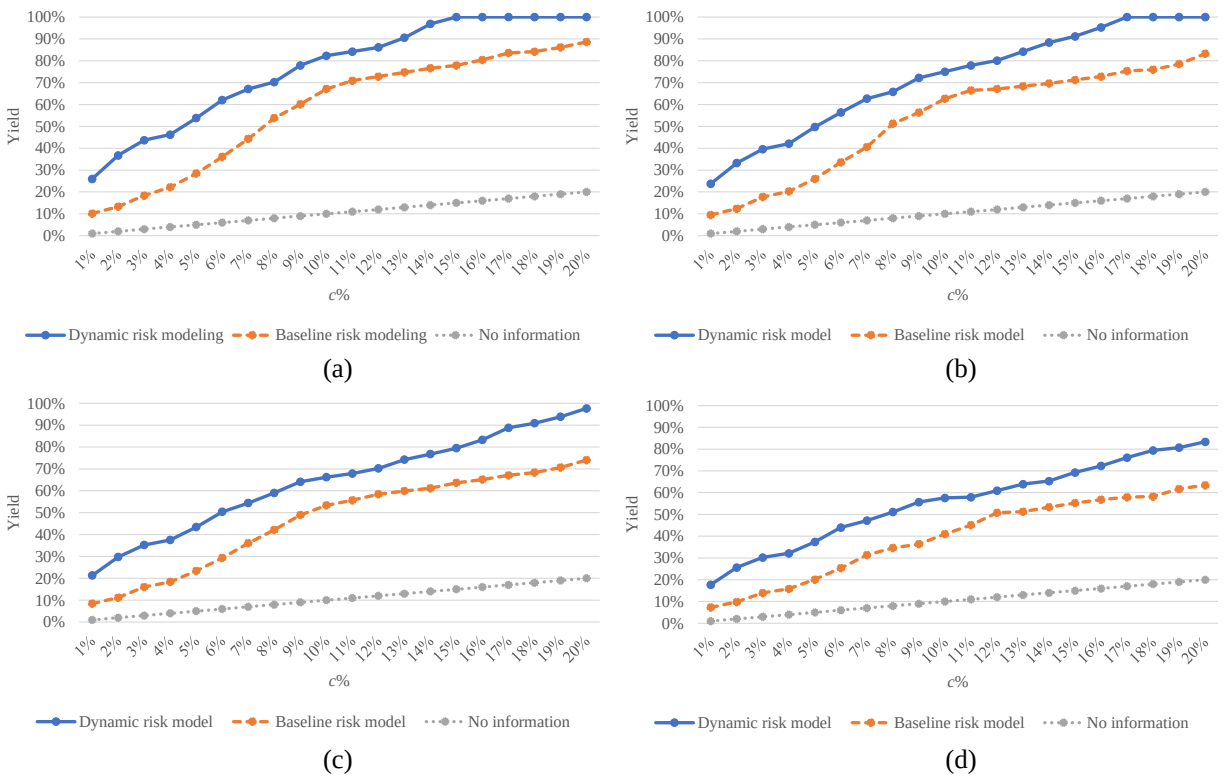
Table 15 Simulation Parameters

Parameter	Specification
Probability of becoming adherent again	Follows the proportional hazards model from Study A
Probability of readmission	Follows the dynamic readmission risk-scoring model
Random effect distribution for adherence model	$N(0,0.269)$
Readmission length-of-stay (LOS) distribution	P(LOS = 1 day) = 0.792, P(LOS = 2 days) = 0.105, P(LOS = 3 days) = 0.065, P(LOS = 4 days) = 0.021, P(LOS = 5 days) = 0.011, P(LOS = 6 days) = 0.006

Table 16 Simulation Results for Strategy B When *n* Is 85 and 90

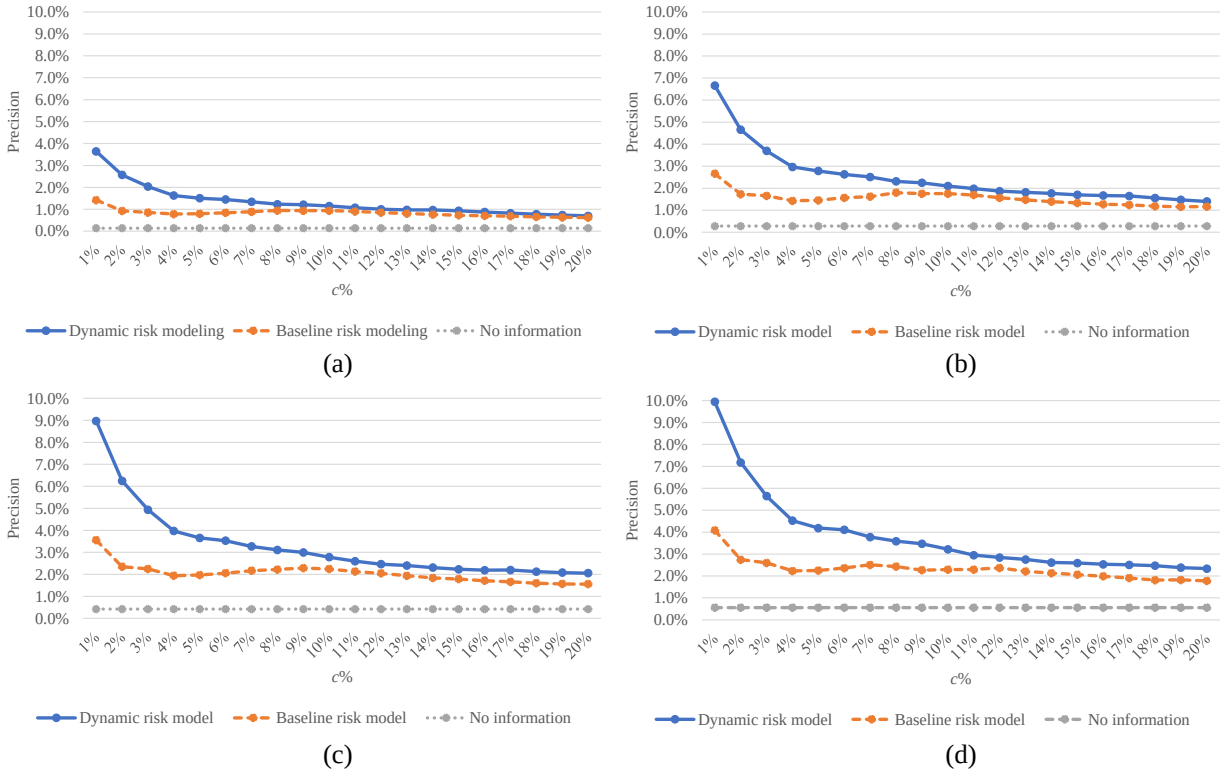
	Scenario	
Outcome	Strategy B ($n = 85$)	Strategy B ($n = 90$)
Total number of manual interventions	831.3	607.3
Call to patient	766.7	559.3
Manual message to patient	58.5	42.9
Call to partner	6.2	5.1
Average individual adherence rate	90.11%	88.69%
Total number of readmissions	289.0	295.7
Number of patients readmitted at least once	262.9	265.7

Figure 9 **Yield Obtained Using Different Readmission Intervention Approaches: Varying the Definition of Patient-Day at Risk**



Note. The definition of *patient-day at risk* in each panel is: (a) patient-day with readmission, (b) patient-day with readmission or prior to readmission (the default), (c) patient-day with readmission or up to two days prior to readmission, and (d) patient-day with readmission or up to three days prior to readmission.

Figure 10 **Precision Obtained Using Different Readmission Intervention Approaches: Varying the Definition of Patient-Day at Risk**



Note. The definition of *patient-day at risk* in each panel is: (a) patient-day with readmission, (b) patient-day with readmission or prior to readmission (the default), (c) patient-day with readmission or up to two days prior to readmission, and (d) patient-day with readmission or up to three days prior to readmission.

Appendix B: Study B – Instrumental Variable Analysis

As we explain in Section 5.3, there may exist unobservable time-varying factors that influence readmission and adherence behavior in addition to time-invariant factors. To address this, we use the instrumental variable (IV) approach to control for both time-invariant and time-varying unobservable factors that may be confounders.

There exist two groups of patients in our study: (1) patients who were usually able to maintain relatively high adherence—they tended to receive few or no interventions when their adherence was high and receive more interventions when their adherence was low; and (2) patients who normally had relatively low adherence and needed intervention to stay adherent—they tended to have higher adherence when they received more interventions and lower adherence when they received fewer or no interventions. For the latter group, we see the possibility of using intervention as IVs for medication adherence. The reasons why we may be able to use intervention as IVs for these patients are because (1) intervention is positively correlated with medication adherence and (2) intervention is likely uncorrelated with unobservable factors that affect readmission. Because of this observation, we focus on this cohort of 115 patients in our IV analysis. Among these patients, the average short- and long- term adherence are 0.512 and 0.544, respectively.

There exist two potentially endogenous binary variables, short-term adherence and long-term adherence. We need at least one IV for each endogenous variable. We want our IVs to be (1) correlated with medication adherence (relevance condition) and (2) uncorrelated with unobservable factors that affect readmission (exogeneity condition). We propose using the number of manual interventions the patient received prior to the present day as IVs. Specifically, for short-term adherence, the IVs that we propose are (1) the number of manual messages the patient received in the past 8 days (i.e., length of adherence window plus five days), (2) the number of phone calls the patient received in the past 8 days, and (3) the number of phone calls the patient's partner received in the past 8 days. For long-term adherence, the IVs that we propose are (1) the number of manual messages the patient received in the past 125 days (i.e., length of adherence window plus five days), (2) the number of phone calls the patient received in the past 125 days, and (3) the number of phone calls the patient's partner received in the past 125 days. By using intervention as IVs, we obtain the effect of medication adherence that varies depending only on how much intervention the patient receives. We will also validate the relevance and exogeneity conditions for these IVs.

Since both readmission and medication adherence are binary variables, we use a bivariate probit model jointly estimated via Full Maximum Likelihood Estimation (FMLE) (Cameron and Trivedi 1998, Greene 2012). We do not include patient fixed effects because patient fixed effects (or individual fixed effects in general) in a two-stage binary response model may cause an incidental parameter problem and biased estimates (Greene 2004). However, since the IV approach already accounts for both time-invariant and time-varying unobservables that may be confounders, it is not necessary to include patient fixed effects. We specify our model for patient i and day t as follows:

$$
\begin{aligned}
ShortTermAdherence^{*}_{it} &= Z_{it}\omega + \theta_1 ManualMessages8Days_{it} + \theta_2 PatientCalls8Days_{it} \\
&\quad + \theta_3 PartnerCalls8Days_{it} + \varepsilon^{1}_{it}, \\
ShortTermAdherence_{it} &= \mathbf{1}\left\{ShortTermAdherence^{*}_{it} > 0\right\}, \\
LongTermAdherence^{*}_{it} &= Z_{it}\Omega + \Theta_1 ManualMessages125Days_{it} + \Theta_2 PatientCalls125Days_{it} \\
&\quad + \Theta_3 PartnerCalls125Days_{it} + \varepsilon^{2}_{it},
\end{aligned}
$$

$$
\begin{aligned}
LongTermAdherence_{it} &= \mathbf{1}\left\{LongTermAdherence^{*}_{it} > 0\right\}, \\
Readmitted^{*}_{it} &= Z_{it}\delta + \beta_1 ShortTermAdherence_{it} + \beta_2 LongTermAdherence_{it} + \alpha_t + \xi_{it}, \\
Readmitted_{it} &= \mathbf{1}\left\{Readmitted^{*}_{it} > 0\right\}
\end{aligned} \tag{8}
$$

where $ShortTermAdherence_{it}$ and $LongTermAdherence_{it}$ are short- and long-term adherence as defined in Study A; $readmitted_{it}$ is a binary variable that equals one when readmission occurs; α_t is fixed effects of days since last discharge; and Z_{it} is a vector of control variables, which include (1) patient characteristics (age, gender, PHQ-2 score, Elixhauser score, and Medicare enrollment), (2) a binary variable indicating weekend and holiday, (3) the number of previous readmissions, (4) days since enrollment, and (5) EA fixed effects.

In Equation 8, the error terms (ξ_{it}, ε^1_{it}) may be correlated to model the endogeneity between the short-term medication adherence and readmission. Similarly, the error terms (ξ_{it}, ε^2_{it}) may be correlated to model the endogeneity between the long-term medication adherence and readmission. We assume that (ξ_{it}, ε^1_{it}) and (ξ_{it}, ε^2_{it}) follow a Standard Bivariate Normal distribution with correlation coefficient ρ_1 and ρ_2, respectively. We can conduct the likelihood ratio tests of null $\rho_1 = 0$ and null $\rho_2 = 0$ to test the presence of endogeneity.

We report the first-stage results in Tables 17 and 18. We find that the effects of the numbers of interventions received on short- and long-term medication adherence are positive and statistically significant. One exception is the effect of the number of manual messages received in the past 8 days on short-term adherence, which is not significant at a 95% confidence level. However, the insignificance is not an issue since we only need at least one IV to be "relevant" to each endogenous variable. The first-stage results suggest that the relevance condition is satisfied.

Table 17 Estimated Effect of Instrumental Variables on Short-Term Medication Adherence

Instrumental Variable	θ (SE)
$ManualMessages8Days$	0.22 (0.12)
$PatientCalls8Days$	0.26^{***} (0.05)
$PartnerCalls8Days$	0.75^{***} (0.06)

Note. Standard error in parentheses. $^{*}(p < 5\%)$,$^{**}(p < 1\%)$,$^{***}(p < 0.1\%)$.

Table 18 Estimated Effect of Instrumental Variables on Long-Term Medication Adherence

Instrumental Variable	Θ (SE)
$ManualMessages125Days$	0.11^{**} (0.05)
$PatientCalls125Days$	0.12^{***} (0.02)
$PartnerCalls125Days$	0.44^{***} (0.02)

Note. Standard error in parentheses. $^{*}(p < 5\%)$,$^{**}(p < 1\%)$,$^{***}(p < 0.1\%)$.

We show the second-stage results in the column titled "With IVs" in Table 19. Similar to the results obtained from the fixed-effects logit model, we find that the effect of long-term adherence is statistically significant while the effect of short-term adherence is not. For long-term adherence, the estimated coefficient of -0.24 translates to an average marginal effect of -0.34%. This marginal effect is slightly smaller than the average marginal effect we obtain using the fixed-effects logit model, which controls for only time-invariant confounders. Moreover, to determine what the results will be if we ignore the potential endogeneity, we run only the second-stage model and report the results in the column

titled "Without IVs". We find that the effects of adherence on readmission are exaggerated when we do not account for either time-invariant or time-varying confounders. Our findings suggest that patients who are healthier and are less likely to be readmitted are possibly more adherent to their medications. Table 20 shows the results of endogeneity tests, which suggest that the error terms in the first- and second-stage equations are negatively correlated. Therefore, endogeneity is present.

Table 19 Estimated Effect of Medication Adherence on Readmission

	β (SE)	
Adherence Measure	With IVs	Without IVs
$ShortTermAdherence$	-0.12 (0.09)	-0.13 (0.10)
$LongTermAdherence$	-0.24** (0.11)	-0.33*** (0.04)

Note. Standard error in parentheses. $^{*}(p < 5\%),^{**}(p < 1\%),^{***}(p < 0.1\%)$.

Table 20 Results of Endogeneity Tests

Parameter	Value (SE)	Test $\rho = 0$ p-value
ρ_1	-0.16 (0.06)	0.04
ρ_2	-0.46 (0.09)	0.00

Note. Standard error in parentheses.

Although we cannot directly verify the exogeneity condition using the bivariate probit model, we can verify the condition using a two-stage least squares (2SLS) model by conducting a Hansen J test of overidenfying restrictions. Although the 2SLS model linearizes binary responses, many economists including Angrist and Pischke (2013) have proposed using 2SLS for nonlinear models with endogenous regressors. Many researchers also adopt a linearized two-stage model partly because it can handle individual fixed effects (e.g., Bavafa et al. 2018). We conduct our analysis using a 2SLS model and find that the results are consistent with those from the bivariate probit model. Using a Hausman's specification test, we find that endogeneity is present. Moreover, to verify the exogeneity condition, we conduct the test of overidentifying restrictions, which is possible because the number of IVs exceeds the number of endogenous variables. Based on the test, we find that the Hansen J statistic is 2.624 with a p-value of 0.623. This suggests that the IVs are uncorrelated with unobservable factors that affect readmission and, therefore, the exogeneity condition is satisfied.