CONDUCTING A RETROSPECTIVE COHORT TO EVALUATE THE ACCURACY OF A REAL-TIME MULTIPLE LOGISTIC REGRESSION TOOL FOR PREDICTING THE LIKELIHOOD OF PATIENT READMISSION

by

Cory F. Hayes

BS, University of Pittsburgh, 2017

Submitted to the Graduate Faculty of

Infectious Diseases and Microbiology

Graduate School of Public Health in partial fulfillment

Of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Cory F. Hayes

It was defended on

December 12, 2018

and approved by

Catherine Haggerty, PhD
Associate Professor, Epidemiology
Graduate School of Public Health, University of Pittsburgh

Mohamed Yassin, MD, PhD
Director of Infection Control
University of Pittsburgh Medical Center – Mercy Hospital

Thesis Advisor:
Jeremy Martinson, DPhil
Assistant Professor, Infectious Diseases and Microbiology
Graduate School of Public Health, University of Pittsburgh

Jeremy Martinson, DPhil

CONDUCTING A RETROSPECTIVE COHORT TO EVALUATE THE ACCURACY OF A REAL-TIME MULTIPLE LOGISTIC REGRESSION TOOL FOR PREDICTING THE LIKELIHOOD OF PATIENT READMISSION

Cory F. Hayes, MPH

University of Pittsburgh, 2018

ABSTRACT

Introduction:

Hospital readmissions have a great deal of public health significance, as they are burdensome and costly to providers, hospitals, and patients. The quality-of-care provided by hospitals is evaluated by comparing hospital readmission rates to national averages. Underperforming hospitals are penalized by the Centers for Medicare & Medicaid Services (CMS). In 2018, 2,597 hospitals are being penalized a total of $564 million. Published studies have demonstrated a multitude of approaches that were successful in reducing readmission rates, but they are too expensive for systemic implementation within a hospital. The University of Pittsburgh Medical Center (UPMC) Mercy Hospital Clinical Analytics team has constructed a multiple logistic regression prediction model that scores patient risk factors in order to flag high risk patients who are most likely to experience readmission.

Objectives/Aims:

Primarily, this study aims to evaluate the accuracy with which the UPMC Mercy multiple logistic regression model correctly predicts readmission risk in a clinical application. Once validated, it is our secondary aim to initiate a discharge intervention specifically for patients who are flagged by the model.

Methods:

The predictive logistic regression tool has been in use for slightly over one year at UPMC Mercy; daily reports score patients as 'lowest', 'lower', 'medium', 'higher', or 'highest' risk for readmission. Based on sample size calculations, about 200 individuals per predicted risk group were retrospectively recruited and followed for the next month in order measure readmission status over 7 days and 30 days. Chi-Square testing for independence and stratified one-sample proportion testing allowed for validation of the model's accuracy.

Results:

Chi-square testing of independence demonstrated that not all risk quintiles had distinct mean readmission rates, contrary to our hypothesis. One-sample proportion testing further illustrated the poor fit for predicting 7-day readmission, with only 1/5 risk groups following expected mean rates of readmission. However, one-sample proportion analysis for 30-day readmission prediction resulted in 3/5 of the risk strata exhibiting similar mean rates of readmission as was expected, as well as a clinically relevant increasing trend across risk strata.

Discussion:

This multiple logistic regression model is not an accurate predictor of 7-day readmission. However, it appears that the model could be clinically relevant for predicting 30-day readmissions. The 'highest' risk strata displayed 36% 30-day readmission, which is 16-18% higher than the national average 30-day readmission rate. Though the model could benefit from optimization, it likely could be utilized in its current state to target high risk groups for 30-day readmission rather reliably.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank UPMC for allowing me the opportunity to research, and for permitting me to utilize this data for my thesis. I would like to especially thank Dr. Mohamed Yassin for agreeing to mentor me in the midst of a busy summer in which his time was already stretched incredibly thin. Every conversation that we shared added some component to this project and strengthened the magnitude of its findings. Finally, I would like to thank my advisor and chair Dr. Jeremy Martinson for not discouraging me from entering this process, despite facing a tight timeline of deadlines. His assistance throughout this process has immensely helped me conquer deadlines.

1.0 INTRODUCTION

1.1 SPECIFIC READMISSION MEASURE SUBTYPES

In the growing body of literature, there is an increasingly diverse repertoire of readmission measures being deployed.[1, 2] It is crucial to clearly define commonly utilized readmission subtypes because they do not all have the same implications in evaluating quality of care[1].

One of the most important factors to consider when evaluating a measure of readmission is the length of follow-up that is measured between discharge from the index admission (the original hospital stay) and the rehospitalization[2]. The most commonly utilized readmission follow-up intervals are 7-day/one week, 14-day/two weeks, 30-day/one month, 60-day/two months, and 90-day/three months[2, 3]. Readmission cases tend to cluster after discharge from the index admission; about ½ occur within 90 days; and about 1/3 occur within 30 days.[2, 3] A longer readmission time interval can detect more instances of readmission than a shorter time interval[2-5]. However, shorter readmission windows are more likely to be directly related to the quality of care received during the index admission[2-5].

Occasionally, studies distinguish planned and unplanned readmission measures[6-9]. This readmission subtype stems from the concept of the planned readmission, in which a readmission is scheduled prior to a procedure[7, 9]. When utilizing this metric, researchers attempt to remove planned readmissions from their dataset, as they are not reflective of care received and are not

preventable[7-9]. For example, one vascular surgery report documented a planned readmission rate of 21.5%, demonstrating that studies overlooking planned readmissions can significantly overattribute readmission rates to poor quality of the surgery[9]. One major challenge of the preventable readmission metric is reliably reaching a consensus on which readmission measures are considered planned, since many studies are retrospective and observational designs[6-9]. As an example, one study focused on preventable readmissions and arbitrarily defined that any readmission connected with obstetrical delivery, chemotherapy, or organ transplant should always be classified as planned, even if a certain date is not planned[7].

Another readmission subtype found widely published within the literature is referred to as a potentially preventable readmission (PPR), also sometimes called a discernable readmission[5, 10]. As the name implies, PPRs are readmissions that are considered to have been unnecessary rehospitalizations that would not have occurred in the presence of proper care, discharge planning, or transitory planning during the index admission[5, 10]. When studying readmissions, researchers concede that there will always be a basal level of readmissions, even under the most ideal circumstances, which are not associated with the index admission[5, 11]. Many researchers strongly advocate for only including PPRs when considering quality of care because they avoid counting the basal readmissions that are not preventable[11]. Relatives of the PPR readmission subtype are the linked and unlinked readmission measures[1, 12]. Linking refers to an association with the index admission; linked readmissions are proposed to be associated with some substandard practice that occurred during the index admission, whereas unlinked readmissions are simply coincidence rehospitalizations with no connection to the index hospitalization[12].

Finally, some studies mention all-cause or hospital-wide readmission rates, whereas others refer to cause-specific or disease-specific readmissions rates[8, 13, 14]. Studying hospital-wide

readmission rates is advantageous because it provides a representation of trends throughout the entire hospital system[13, 14]. Studies that focus on disease-specific readmission benefit from understanding how readmission affects a specific stratum of patients[14-16].

1.2 READMISSION CASE DEFINITION

Unless otherwise specified, every occurrence of 'readmission' within this manuscript will refer to hospital-wide, all-cause readmissions without risk adjustment[8, 17]. This measure of readmissions is broadly inclusive and conservative, including all recorded instances of readmissions regardless of whether they were planned, linked to the index admission, or deemed preventable[8, 17]. The hospital-wide all-cause readmission measure dodges the inconsistencies, discrepancies, and biases that arise when panels must reach a consensus opinion in classifying readmissions are related or not to the index admission[5, 11, 18]. Observational stays that were recorded in medical charts are treated as admissions or readmissions in this study, despite that the CMS does not consider observational stays as admissions[19].

As was previously mentioned, 7-day readmissions and 30-day readmissions can offer different information with regards to quality-of-care; 7-day readmissions are more likely to be linked to index cases, whereas 30-day readmissions provide a follow-up window that tracks most of the patient readmissions that occur[1-4]. Both measures are important to the research questions of interest, so 7-day and 30-day readmissions are both recorded and specified accordingly within this study.

1.3 READMISSIONS AND QUALITY OF CARE

In the 1990's many studies began associating hospital readmissions with potential substandard clinical quality of care[11, 20, 21]. One of the original studies to identify this correlation did so by measuring quality of care with a metric that they titled 'readiness-for discharge'[21]. In this section, clinical staff were scored for how well patients and their families were educated and informed, how clinically stable the patient was upon discharge, and what follow-up medical care was provided post-discharge[21]. Patients with heart failure or diabetes who scored in the lowest quartile in readiness-for-discharge surveys were twice as likely to be readmitted for the same condition compared to those who did not score in the lowest quartile[21]. Two years following the initial study, the same group conducted a meta-analysis of sixteen different clinical quality studies, where quality of care was ranked as substandard, normative, or exceptional[20]. This landmark study revealed that care categorized as relatively low quality increased the odds of readmission by a factor of 55 when compared with care categorized as higher quality[20, 22].

Within the past decade, readmissions have emerged as a standard method of evaluating clinical quality of care[11, 18, 23-28].Unlike many measures, readmission statistics concentrate on patients at their highest vulnerability, who are attempting to transition from inpatient care back to a healthy lifestyle[5]. Though readmissions can arise naturally through unavoidable, natural sequalae of disease or patient frailty, the variability of readmissions between different hospitals indicates that some facilities are operating at a higher standard than others[5, 29, 30]. Programs have identified that strictly implementing aseptic protocols, educating patients prior to discharge, planning transitory care, focusing on medication needs, and communicating clearly with patients decreases readmission rates significantly[7, 16, 28-30]. In a trial studying the effects of nosocomial infections, researchers concluded that surgical site infections increase the rate of 30-day

readmission by about 2-5.5 times compared to patients who did not suffer from hospital acquired infections[22]. Patients who experienced unplanned 30-day readmission were 55% more likely to have experienced poor quality of care during their index admission[22].

1.4 HOSPITAL READMISSION INDUCED BURDEN

Repeated rehospitalizations are burdensome to patient recovery due to placing patients at heightened risk of developing adverse complications[14, 23, 31]. Many patients experience some degree of 'post-hospital syndrome', which is a period of increased vulnerability after discharge from a hospitalization resulting from increased stress, poor sleep, interrupted nutritional intake, and adverse drug effects[14, 31, 32]. A 2011 study of colon cancer patients found that patients who were readmitted within 30 days experienced 16% mortality compared to a 7% rate in those who did not experience 30-day readmission[3]. Another research group directly compared newly admitted acute ischemic stroke (AIS) patients to patients being re-hospitalized for AIS[33]. In this observational study, researchers concluded that readmitted patients had higher mean length of stays (LOS), mortality rates, and medical costs than index cases[33].

Hospital readmissions are incredibly costly to patients, providers, and hospitals. In one study that matched cases of *Clostridium difficile* Associated Diarrhea (*C. diff*) to patients who did not have *C. diff*, researchers concluded that *C. diff* increased patient length of stay by an average of 4.7 days, translating to an additional cost of about $7,286[34]. In another publication associating hospital quality and cost of stay, the authors determined that patients attending hospitals ranking in the highest quintile of complication rates paid an average of $2,400-$5,400 more for their stay[35]. In 2013 estimates, index admissions cost patients and providers about

$13,100, whereas the average cost of readmission was $13,800 in addition to the cost of the index stay[36]. In 2013, the Healthcare Cost and Utilization Project (HCUP) calculated that the average cost of readmission for Medicaid or privately insured patients was about $3,000 more than the cost of the corresponding index admission[36]. In 2011, HCUP estimated $41.3 billion in direct hospital costs associated with 3.3 million adult all-cause 30-day readmission cases[8].

For Medicare covered patients, Heart Failure, pneumonia, and septicemia result in the three most frequent causes of 30-day readmissions, resulting in about 316,000 readmissions and $4.3 billion in hospital costs[8]. In Medicaid covered patients, the three most frequent conditions associated with readmissions are mood disorders, schizophrenia, and diabetes, which resulted in about 101,000 cases of 30-day readmission and about $839 million in hospital costs[8]. In privately insured patients, the three most common 30-day readmission cases were related to maintenance chemotherapy, mood disorders, or surgical or medical complications, resulting in about 63,000 readmissions totaling $785 million in healthcare costs[8]. Finally, there are indirect burdens that readmissions cost patients, such as loss of income, loss of career/future employment, and quality of life factors like spending time away from family. One example is that about 1/5 of readmissions are at a different healthcare facility than the original hospitalization, which often leads to duplicate testing, unnecessary interventions, or delayed treatment and diagnosis[27]. Additionally, indirect impact readmissions can harm employers, who spend $260 billion in work related losses every year[37, 38]. Aside from readmission induced financial burdens, unplanned readmissions also disrupt clinical healthcare systems, unnecessarily diverting resources away from other patients who need medical attention[6].

1.5 HOSPITAL READMISSIONS REDUCTIONS PROGRAM (HRRP)

1.5.1 HRRP Aims

At the end of every fiscal year, the Center for Medicare & Medicaid Services (CMS) financially reimburses hospitals for Medicare and Medicaid insured patients that received care[15]. It is estimated that the CMS spends over $17 billion annually reimbursing hospitals specifically for costs associated with unplanned readmissions, which is 17% of the total CMS reimbursement budget[26, 28]. Of the CMS money spent on readmissions, it is further estimated that $12 billion is wasted on 'potentially preventable' readmissions[16]. Some theorize that since hospitals receive the same amount of money for treating a readmitted patient as they do for treating an index case, readmissions perversely benefit hospitals financially[16].

Given this information, the Affordable Care Act (ACA) was enacted into United States law in 2010, which contained the blueprint for the Hospital Readmissions Reductions Program (HRRP), though it would not take full effect until 2012[15, 19, 24]. The HRRP relies on evaluating hospitals based on readmission rates; a hospital with unusually high readmission rates likely is not performing the same standard of care as a hospital with low readmission rates[15]. The Hospital Readmissions Reductions Program aims to improve patient outcomes, decrease inefficient Medicare expenditures, and establish hospital culpability and transparency[15, 19, 24].

About 20% of Medicare beneficiaries experience 30-day hospital readmission, and about 34% experience 90-day hospital readmission[26]. With regards to Medicare 30-day readmissions, rates range from 14% to 22% in the lowest and highest performing deciles of states respectively; this variation reinforcing the inconsistencies in care quality and opportunity for improvement[14, 39]. Approximations posit that a reduction of only 10% of avoidable readmissions would save the

CMS $1 billion dollars yearly[28]. If all hospitals achieved the levels of the current highest performers, there would be $1.9 billion in annual savings[30, 39]

1.5.2 HRRP Methodology

Because of the basal level of unpreventable readmissions that potentially have nothing to do with the index admission or quality of care received, the HRRP initially only targeted three very frequent cause-specific readmission rates that the Center for Medicare & Medicaid Services (CMS) deemed to have high rates of preventable readmission in 2012: Acute Myocardial Infarction, Heart Failure, and Pneumonia[5, 15]. In the years since the HRRP's inception, yearly revisions have added the following diseases to the cause-specific readmission rates that are monitored: acute exacerbation of chronic obstructive pulmonary disease, elective total hip arthroplasty, elective total knee arthroplasty, and coronary artery bypass graft surgery.

Upon its ratification, the HRRP section of the ACA established that these measures would be measured immediately and published each fiscal year[15]. In theory, by risk adjusting hospital readmission rates and comparing them nationally, underperforming hospitals can launch initiatives to reduce the readmission rates[16]. These readmission rates are also posted publicly in an act of transparency from the CMS so that patients can compare hospitals based on quality of care before choosing one[15].

1.5.3 HRRP Financial Incentivization

The final, and most crucial section of the HRRP is the incentive for hospitals to participate[15, 19, 24]. Beginning in 2013, a hospital's prior three years of unplanned cause-specific 30-day

readmission rates were risk adjusted for factors such as age and sex, and then compared nationally to other hospitals. In 2013, a hospital could be fined a maximum of 1% of its total CMS reimbursement.[15] In 2014 the cap was raised to 2%, and from 2015 onwards it has been capped at 3%[15, 19, 24]. Financial incentives are essential to ensuring cooperation in efforts to improve national clinical quality of care[16, 28].

These calculations rely on a measure referred to as the standardized readmission ratio (SRR), also called the excess readmission ratio (ERR)[15, 19, 24]. The ERR is calculated individually for each cause specific readmission that the CMS chooses to track[19]. Essentially the ERR is a ratio that evaluates how a hospital is performing compared to how similar hospitals performed; any ratio over 1 indicates higher than expected readmission rates for that specific disease[19]. Then, these ERRs are then considered together in the total payment adjustment factor calculation, which determines how much a hospital is penalized (Figure 1)[15, 19, 24].

$$P = 1 - \min\left\{.03, \sum_{dx} \frac{Payment(dx) * \max\{(\mathrm{ERR(dx)} - \mathbf{1.0}), \mathbf{0}\}}{All\ payments}\right\},$$

Figure 1. HRRP Payment Adjustment Calculation

Payment adjustment factors (P) are hospital specific calculations[19]. In Figure 1, dx represents the disease specific readmission cohorts that the CMS monitors (AMI, COPD, HF, Pneumonia, CABG, and THA/TKA)[15, 24]. Payment dx, is the full CMS reimbursement to a hospital for a specific disease, prior to any assessed penalties[15, 24]. We can see that if a hospital has an excessive readmission rate, payment dx is multiplied by the ERR. However, if the ERR ≤ 1, then the disease specific term is dropped from the calculation. When all of the disease specific readmission terms are summed together, the payment adjustment factor is equal to 1 minus the

sum of those terms. The term 'min .03' seen in Figure 1 signifies that if readmission terms add up to greater than .03, P is simply equal to 97%, representing the 3% cap on HRRP hospital fines[15, 19, 24]. Beginning in FY 2019, this formula also will weigh the case mix that a hospital faces, and only compare the ERR to other hospitals with a similar case mix[24]. This new stratification will adjust for socioeconomic status (SES), and ensure that hospitals are penalized for suboptimal quality of care rather than treating lower SES populations who are more vulnerable to readmission[24].

1.6 READMISSION TRENDS SINCE HRRP IMPLEMENTATION

In a 2018 survey, 66% of interviewed hospital administrators believed that HRRP implementation had altered systemic efforts to reduce readmissions within their hospital[40]. However, the best way to evaluate the impact of the HRRP initiative is to observe the change in readmission trends since its implementation. Between 2009-2013, readmission for Medicare patients 65 and older dropped to 16.2% from 17.3%[36]. In 2013, the CMS reported that the all-cause 30-day readmission rate, which had been fixed on 19% for five years, dropped to 17.8%[28].The observed reductions in readmission rates primarily stemmed from improvements in HRRP targeted diseases, which declined sharply within the first two years after the implementation of the HRRP[40-42]. Nonetheless, there was also an observed spillover effect, in which nontargeted readmission rates fell by about 1%[25, 33, 41-43]. Interestingly, Desai et al. describe that no change in readmissions was observed in the two-year window of public reporting prior to the CMS monetary penalization[43]. However, after the fines began, the sharp declines were observed[43].

Of note, there are a few concerns that researchers and healthcare professionals are monitoring. First, a 34% increase of the use of observational stays occurred close in proximity to the onset of the HRRP program[28]. Similarly, in 2010 and 2011, the top 10% of hospitals with the biggest improvement in hospital readmission reduction were discovered to have increased their usage of observational stays in the same period by about 25%[28]. By utilizing observational stays, a patient is not truly admitted, so their return the hospital does not count as readmission against that hospital[28]. Another concern arises when readmission calculations lump surgery in with medical admissions. Unlike with medical hospitalizations, the surgical procedure and unavoidable surgical complications are a major risk factor for readmission[10]. This is demonstrable in that the spillover effect that was discussed previously was not observed in surgery patients[40]. Finally, policymakers are focused on monitoring hospital practices to ensure that readmission reductions are truly related to improvements in patient outcomes, while minimizing unintended consequences[28].

1.7 TRENDS IN HRRP INDUCED HOSPITAL FINES

In the first year that fines were levied against hospitals for excess readmission rates (FY 2013), hospitals were fined a total of $290 million[15]. In FY 2018, that has climbed to 80% of hospitals that are being fined a total of $564 million[15]. Table 1 demonstrates how even a small percentage of reimbursement withholding can impact a large hospital system like UPMC. In FY 2013 and FY 2014, the UPMC hospital system was fined less than $1 million. However, by FY 2017 and FY 2018, UPMC is being fined close to $7 million annually.

Table 1. UPMC Readmission Penalties Since the Introduction of HRRP Penalizations

	FY 2013	FY 2013 %	FY 2014	FY 2014 %	FY 2015	FY 2015 %
Altoona	-$186,715	- 0.36%	-$35,389	- 0.07%	-$495,695	- 0.95%
Bedford	-$1,085	- 0.03%	-$2,499	- 0.07%	$0	+ 0.15%
East	$0	0	$0	0	-$18,440	- 0.10%
Hamot	-$16,139	- 0.03%	-$98,444	- 0.15%	-$210,549	- 0.29%
Horizon	-$25,408	-0.14%	-$9,699	- 0.06%	-$50,273	- 0.29%
Jameson	-$161,831	- 0.82%	-$178,925	- 0.92%	-$176,657	- 0.92%
Kane	-$19,833	- 0.50%	$0	+ 0.12%	$15,446	+ 0.49%
Magee-Women's	-$23,731	- 0.74%	-$23,088	- 0.59%	-$7,510	- 0.22%
Mckeesport	-$56,526	- 0.28%	$0	+ 0.01%	-$48,057	- 0.26%
Mercy	-$40,468	- 0.06%	-$72,143	- 0.12%	-$308,061	- 0.50%
Northwest	-$20,196	- 0.09%	-$86,639	- 0.43%	-$109,158	- 0.55%
Passavant	-$107,749	- 0.27%	$0	+ 0.07%	-$482,118	- 1.28%
Presby/Shadyside	$0	+ 0.08%	$0	+ 0.08%	-$2,763,869	- 0.97%
St. Margaret	-$45,258	- 0.13%	-$5,996	- 0.02%	-$271,531	- 0.73%
UPMC Fine Total	-$704,939		-$512,822		-$4,926,472	

	FY 2016	FY 2016%	FY 2017	FY 2017%	FY 2018	FY 2018%
Altoona	-$452,906	- 0.75%	-$666,357	- 1.11%	-$1,000,873	- 1.67%
Bedford	-$2,076	- 0.06%	-$28,740	- 0.85%	-$2,685	- 0.08%
East	-$10,247	- 0.05%	-$91,020	- 0.48%	-$13,646	- 0.07%
Hamot	-$383,651	- 0.54%	-$1,460,444	- 2.07%	-$545,942	- 0.77%
Horizon	-$3,811	- 0.02%	$0	+ 0.40%	-$16,568	- 0.10%
Jameson	-$64,055	- 0.41%	-$98,519	- 0.63%	-$18,947	- 0.12%
Kane	-$13,983	- 0.55%	$0	$0	$0	+ 0.30%
Magee-Women's	-$38,649	- 1.30%	-$2,574	- 0.09%	-$10,817	- 0.37%
Mckeesport	-$62,936	- 0.37%	-$13,988	- 0.08%	-$23,991	- 0.14%
Mercy	-$284,275	- 0.53%	-$294,097	- 0.55%	-$244,471	- 0.46%
Northwest	-$15,481	- 0.08%	-$97,976	- 0.53%	$0	+ 0.08%
Passavant	-$238,635	- 0.60%	-$623,777	- 1.56%	-$386,192	- 0.96%
Presby/Shadyside	-$3,007,909	- 1.15%	-$3,550,821	- 1.36%	-$4,058,910	- 1.56%
St. Margaret	-$13,420	-0.04%	-$363,042	- 1.10%	-$409,422	- 1.25%
UPMC Fine Total	-$4,592,034		-$7,291,355		-$6,732,464	

1.8 INTERVENTIONS TO REDUCE READMISSION RATES

There are a plethora of successful programs that have reduced hospital readmission rates in the body of literature. These interventions generally fall into three broad categories: pre-discharge interventions, post-discharge interventions, and bridge interventions that act before and after discharge to bridge the two[44]. Successful pre-discharge interventions tend to focus on standardized discharge procedures that act like a checklist to ensure a patient is receiving all pertinent information[14, 18, 26, 30, 45].Other pre-discharge interventions can include improving nursing-to patient-ratios, focusing on patient understanding and education, or improving clear communication to the patient[14, 46]. Transitory interventions concentrate on providing outpatient care, education of adverse symptom monitoring, and patient resources external to the hospital[18, 33, 47, 48]. Telephone check-in calls are a very simple and common successful post-discharge intervention that is found widespread throughout the literature[46, 49].

1.9 PREDICTIVE MODELING TO LOWER READMISSION RATES

Despite the variety of proven intervention strategies that exists, to implement them a hospital must sacrifice a large, upfront financial investment in the payment of additional personnel, cost of training programs, and the coordination of care[6, 13, 35, 50-52]. Hospital administrators are faced with the task of minimizing readmissions and associated HRRP penalties whilst simultaneously maximizing patient outcomes and saving money[13, 44, 53, 54]. Predictive modeling allows researchers to identify high-risk populations and focus interventions on a specific stratum of patients rather than implementing them systemically[1, 13, 55].

There are many patient risk factors that have been statistically associated with heightened risks of hospital readmissions. The most frequent risk factors in readmission models are age, gender, number of comorbidities, prior hospital/ED visits, Medicare insurance, and LOS[4, 18, 47, 50, 52, 54, 56, 57]. Primary illness is a unique risk factor that can be utilized to estimate likelihood of readmissions; a 2013 Healthcare Cost and Utilization Project (HCUP) report on readmissions determined that the five highest rates of 30-day disease-specific readmissions were associated with sickle cell anemia (31.9%), gangrene (31.6%), hepatitis (30.9%), diseases of the white blood cells (30.9%), and chronic renal failure (27.4%)[58]. Aside from disease specific traits, SES is a commonly included risk factor in predictive readmissions modeling[50]. Patient labs were another frequently utilized risk factor in modeling, which included things such as hemoglobin counts and white blood cell counts[50]. Mental health conditions also were included in many of the readmission models, specifically including depression, anxiety, schizophrenia, and Bipolar Disorder among others[13, 53, 59].

Many predictive models for hospital readmission have been synthesized for the exact purpose of targeting high risk individuals for treatment. Kansangara et al., 2011 reviewed 26 unique predictive readmission models from various publications, and concluded that all models poorly predicted high risk patients[18]. Zhou et al., 2016 then furthered pursued this study with the same analysis of predictive readmission models that were published after the Kansangara et al. review. After reviewing 73 models, the Zhou et al. concluded that only two models, both based on potentially preventable readmissions, achieved high discriminative ability in their predictions[50].

One of the key contributions from the Kansangara et al. systematic review is a list of criteria that the authors identify as the necessary components of an ideal predictive model for

readmissions. They describe that an optimal clinical readmission predictive model should be able to provide data prior to healthcare administrators and physicians prior to discharge, should reliably discriminate high-risk and low-risk patients, and should be adapted to the specific setting in which it is used[18].

2.0 PROJECT AIMS

Hospital readmissions are burdensome and costly to providers, hospitals, and patients; financial penalties and loss of beds and staff damage hospitals, loss of income and poorer prognoses impact patients, and the excess costs of hospital stays harm insurers. Given that no party benefits from readmissions, there have been numerous efforts to reduce them. However, the major setback in reducing readmission is determining how the interventions are employed within healthcare systems. Because hospitals have finite resources, interventions must focus on certain risk groups, comorbidities, hospital units, etc.

To address this issue within the UPMC hospital system, UPMC clinical analytics team analyzed past patient data to synthesize a real-time logistic regression model that weighs patient traits, such as age, gender, whether a patient had experienced an admission in the past year, etc. The model then considers these patient traits and scores patients in risk quintiles ranging from 'lowest' risk to 'highest' risk for future readmission. The ultimate goal of the prediction tool is to alert healthcare workers to which patients are the most likely to be readmitted. Nurses and physicians can then flag high risk patients for strategic and cost-effective discharge interventions. Reports with patients' predicted readmission scores are compiled daily and emailed to physicians at UPMC Mercy. This system has been active within UPMC Mercy for slightly over one year. Unfortunately, the validity of the model's predictions remains unknown, so it is not utilized routinely as a diagnostic tool like it was intended to be.

Our primary aim within this study is to analyze the accuracy of this clinical prediction model and to validate this instrument for reliable usage in clinical care. In the primary step, we aim to ensure that individuals across risk quintiles have distinct patterns of readmission. We hypothesize that the five risk quintiles to which patients are assigned will exhibit statistically different mean 7 and 30-day readmission rates from each other. Second, we aim to compare the model's predicted mean readmission rates to the observed mean readmission rates for each assigned risk group within our cohort. We hypothesize that none of the five assigned risk groups will have significantly different observed and expected 7-day or 30-day mean readmission rates.

3.0 METHODS

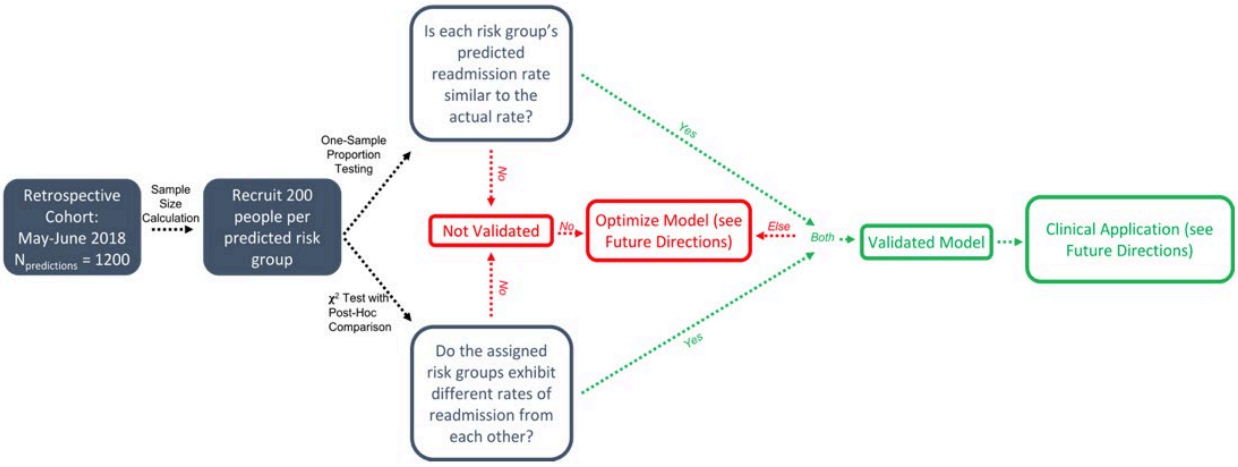


Figure 2. Experimental Design Flowchart

3.1 MULTIPLE LOGISTIC REGRESSION

3.1.1 Introduction to Logistic Regression

A. $$Ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$$

B. $$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Figure 3. Logistic Regression Calculations

Multiple logistic regression is an integral analytic technique, used to estimate the probability that a binary event will occur based on weighing multiple covariates or risk factors. Figure 3 details the various formulae that are associated with logistic regression. Figure 3A displays how the log probability ($log\frac{p}{(1-p)}$), referred to as 'logitP', is calculated. β_0 is the y intercept, and broadly can be interpreted as the log odds of the binary outcome when none of the risk factors are present, though exceptions to this interpretation arise when a risk factor is continuous (i.e. age of 0 isn't really possible). There are k variables/risk factors that are included in a multiple logistic regression model. X_j represents each variable value that is factored into the model, and β_j is the associated coefficient for each variable. Based on all of the terms described thus far, Figure 3A for LogitP can be rewritten more simply as $\beta_0 + \sum_{i=1}^{k} \beta_j x_j$ in which each individual risk factor is associated with its own $\beta_j x_j$ term. The sum of all k risk factor terms and

the y intercept then provides the estimated logitP. Once all of the risk terms and the intercept are summed, the logitP can be transformed into a probability through the equation detailed in Figure 3B. If the logitP value is calculated in part 1A, it can be substituted into exponential term that's in Figure 1B, such that: $p(x) = \frac{1}{1+(e)^{-logitP}}$.

3.1.2 Accounting for Various Variable Types

X_j can take on a different meaning based on the type of variable that x codes for. When considering nominal categorical variables, such whether a patient has diabetes (Y/N), it is standard practice to code '1' for yes and '0' for no. If the patient has diabetes, the term $\beta_{Diabetes} * (1)$ is included in the model, and the term will equal the beta coefficient for diabetes. However, if the patient does not have diabetes, the term $\beta_{Diabetes} * (0)$ would be included in the model, cancelling out the diabetes term in this calculation. When x is a continuous variable, such as age, x simply will equal the patient's age. If a patient's age is 50 years old, the age term will be $\beta_{Age} * (50)$. Finally, there are situations in which ordinal variables are considered, like with systolic blood pressure ranges:

<120mmHg = normal

120-129mmHg = elevated,

130-139mmHg = stage 1 hypertension,

>140mmHg = stage 2 hypertension.

It would be standard practice in this example to code normal blood pressure as '0', elevated as '1', stage 1 hypertension as '2', and 'stage 2 hypertension as '3'. When ($\beta_{Systolic} x_{systolic}$) is then calculated for a patient who falls within normal range of blood pressure, (0 * $\beta_{Systolic}$) drops the term from the model. When $x_{systolic}$ takes on the other potential values (In the blood pressure example 1,2, or 3), $\beta_{Systolic}$ simply compounds by each additional level of ordinality.

3.1.3 Creation of the Readmission Multiple Logistic Regression Model

Within the context of the current study, the binary outcome of interest is whether or not a patient is readmitted during the follow-up time of 7 or 30 days. The UPMC clinical analytics team initially considered over 100 covariates within univariate analyses that were based on forest plots synthesized from systematic review of readmission literature. These covariates were evaluated for significance within a training dataset of 1,000,000 prior UPMC patients. Variables that were statistically significant within the univariate analyses were then considered for multiple logistic regression. Covariates found to be significant in the multiple logistic regression analysis then were analyzed for multicollinearity, and were combined or dropped from the model if covariates were significantly related with each other. This final model had strong specificity, sensitivity, and c-statistic scores.

Unfortunately, many of the covariates considered in this regression model were extracted from patient labs, diagnostic factors, and complications that were not added to the patient's chart for at least 3 days post-discharge. For example, the literature states that about 30% of patients

admitted for Sickle Cell Anemia are readmitted within 30 days, and the initial regression model found that it was a very reliable predictor of readmission[58]. However, blood cultures are not analyzed and added to the EPIC electronic medical records until at least 3 days post-admission, and sometimes one-week post-admission. Though many of these covariates that were dropped increased the validity and reliability of the logistic regression model's predictions, the model itself is not clinically useful if scores cannot be obtained until after a patient is discharged. Due to this, the UPMC clinical analytics team dropped all of the covariates that would not be available during the first 24 hours of a patient's index admission from the final multiple logistic regression model.

3.2 SAMPLE SIZE CALCULATIONS

Following standard practices, it was determined that our type I error would be held fixed at 5%. In order to minimize the risk of type II error, the number of participants needed per risk group was calculated based on 80% power. Because sample size calculations are dependent on the primary outcome analysis, the aforementioned calculations were based on the one-sample proportion testing. Single proportion sample size calculations were based on the formula:

$$n \geq \frac{z_{\alpha/2}}{M} * \hat{p}(1-\hat{p})$$

n = the number of people needed per group

M = Margin of error

$\hat{p}$ = Estimated probability of readmission

However, the expected proportion of readmissions is not consistent across risk strata. Because the $\hat{p}(1-\hat{p})$ term is mathematically maximized with a value of 0.5, all sample size calculations conservatively set the likelihood of readmission under the null hypothesis to 50%.

3.3 STUDY DESIGN AND ENROLLMENT

At the onset of this study, the logistic regression tool had predicted patient readmission risks at UPMC Mercy for slightly over one year. This type of data lends itself to a retrospective cohort design. Because the model's past predictions were made in real-time and recorded in a spreadsheet, the quality of the predictions could be evaluated by searching those patients' Electronic Medical Records (EMRs) to determine if patients were actually readmitted within 7 or 30-days. Figure 4 details both the study design and the timeline.

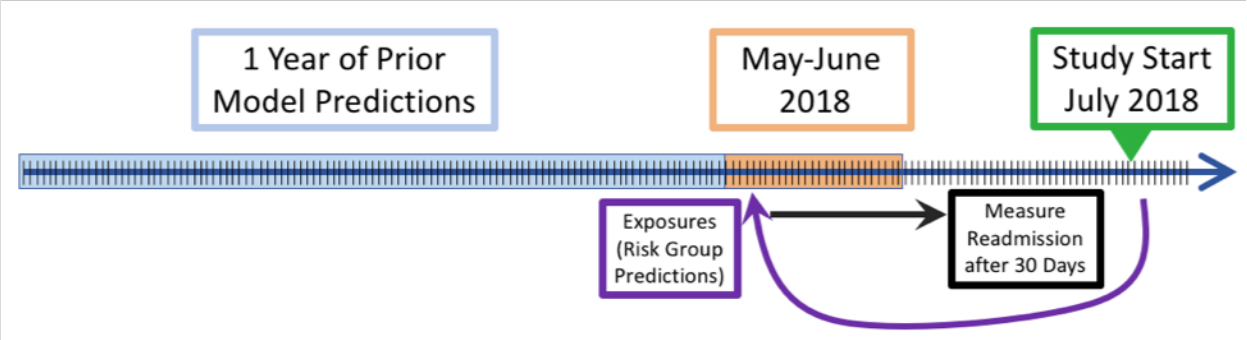


Figure 4. Retrospective Cohort Design and Timeline

Based on the sample size calculations, it was determined that about one month of logistic regression predictions would suffice to ensure the appropriate study power. Because the study

began in July 2018, predictions from the most recent month with a full 30-days of follow-up time were recruited and analyzed.

3.4 STATISTICAL ANALYSES

All analyses were conducted with a fixed alpha of .05 in order to minimize the risk of Type I error. Calculated p-values were then directly compared to this fixed alpha value to determine significance. All statistical tests were calculated using Stata SE 14 or SAS 9.4.

3.4.1 Demographic Analysis

Demographic analysis was necessary to ensure that the recruited sample population did not significantly differ from the 25,000-person training set. It is worth noting that since age and sex are weighted in the model, they cannot be adjusted for if significant differences exist in the testing population and the actual patient population. Nonetheless, it is absolutely crucial to carry out demographic analysis to ensure that the training set mirrors the demographics of the patients in this study. Otherwise the model's coefficients will be tailored to a different population, and the predictions will not be accurate. Age, sex, and length of stay were compared between predicted risk groups. All continuous means were tested for significance with ANOVA, and all categorical hypothesis testing relied on chi-square analysis.

3.4.2 Chi-Square Analysis

Chi-square testing is essential in determining that the predicted risk groups actually experience distinct mean proportions of readmission. Initial chi-square analysis tested the alternative hypothesis that at least one risk group was distinct from the others:

H_0: $p_{lowest} = p_{lower} = p_{medium} = p_{higher} = p_{highest}$

H_1: At least one risk group's mean proportion of readmission is significantly different from

the other risk groups' mean proportions of readmission.

Testing were conducted independently for 7-day and 30-day readmission. The initial chi-square tests were followed by post-hoc testing, utilizing pairwise comparisons to identify which risk groups were and were not significantly distinct from each other. Because pairwise tests rely on multiple comparisons, adjustments must be made to make α more conservative. While Bonferroni tests are often a strong adjustment tool for multiple comparisons, it is not an appropriate technique in this scenario. Bonferroni adjustments divide the α value by the number of pairwise comparisons. Because our analysis includes all ten pairwise comparisons, the alpha value considered for rejection would be 0.005, which is overly conservative in this context. A Tukey's adjustment for multiple comparisons is a more appropriate statistical technique for this analysis, because all pairwise tests are considered in a single step and control for standard error between mean proportion rates. All crude p-values and Tukey adjusted p-values are provided in the χ^2 results section so that the effect of the adjustment is visible.

3.4.3 *One-Sample Proportion Testing*

When the multiple logistic regression model was applied to the 25,000-person dataset of past UPMC Mercy patients, mean proportion rates of 7 and 30-day readmission were calculated for each stratified risk group. Each risk group's mean proportion of readmission was utilized as an expected rate in a one-sample proportion test. Hypothesis testing was conducted at 7-days and 30-days:

H_0: $p_{Kobs} = p_{Kexp}$

H_1: $p_{Kobs} \neq p_{Kexp}$

In this testing, k designates the kth risk group. In this study, k will be one of the following: lowest, lower, medium, higher, or highest. Obs represents the observed mean proportion for the kth stratum in our retrospective cohort, while exp designates the expected proportion of readmission based on the training dataset. Testing concluded whether or not observed stratum specific mean proportion rates of 7 and 30-day readmission significantly differed from the expected stratum specific mean proportion readmission rates.

4.0 RESULTS

4.1 LOGISTIC REGRESSION MODEL

Patient predictions for likelihood of readmission were calculated from covariate coefficients and p-values that are listed in Figure 5A. As is observable from Figure 5A, a patient having experienced a hospital admission within the previous year is by far the largest predictor of readmission within the context of this model; so much so that there are two separate equations with different coefficients depending on whether a patient was or was not readmitted within the past year. Aside from prior admissions, low SES population (Lpop: yes/no) was another covariate that explained a large proportion of readmission predictions. If not admitted within the past year, patients who are from low SES population experience 2.06 times greater odds of experiencing readmission compared to those who are not low SES population and were not readmitted in the past year. If patients were admitted within the past year, being of low SES makes them 1.55 times more likely to be readmitted compared to those who are not low SES. Other significant risk factors included within the model were age, gender, whether a patient was transferred from a skilled nursing facility (SNF), whether a patient was a transfer from another hospital, if a patient had visited the emergency department within the past year, and if the patient carried Medicare insurance (commercial rollup). The category titled Transformed Previous

Hospital Visit, indicates that the readmission is directly related to the prior hospital visit, which is why it is only applicable to the model in which a prior hospital visit had occurred.

Because patient charts are maintained in an electronic medical record (EMR), X_J patient risk factors from Figure 5A are automatically collected from patient charts via the logistic regression tool. In this sense, the tool creates patient readmission risk scores in real-time based on passively collected data from patient charts. Instead of directly being quantified as a probability of readmission, the calculated logitP probability is translated into categorical risk quintiles: Lowest, Lower, Medium, Higher, and Highest risks. In order to determine the cutoffs for the quintiles, data researchers on the UPMC Clinical Analytics team rigorously conducted and calculated analyses of sensitivity, specificity, ROC curves, c statistics, and confusion matrices to maximize accuracy of the model. Of particular note, this team trained the model to more accurately predict readmissions in the highest and lowest quintiles, since the extremes were of highest interest for clinical application.

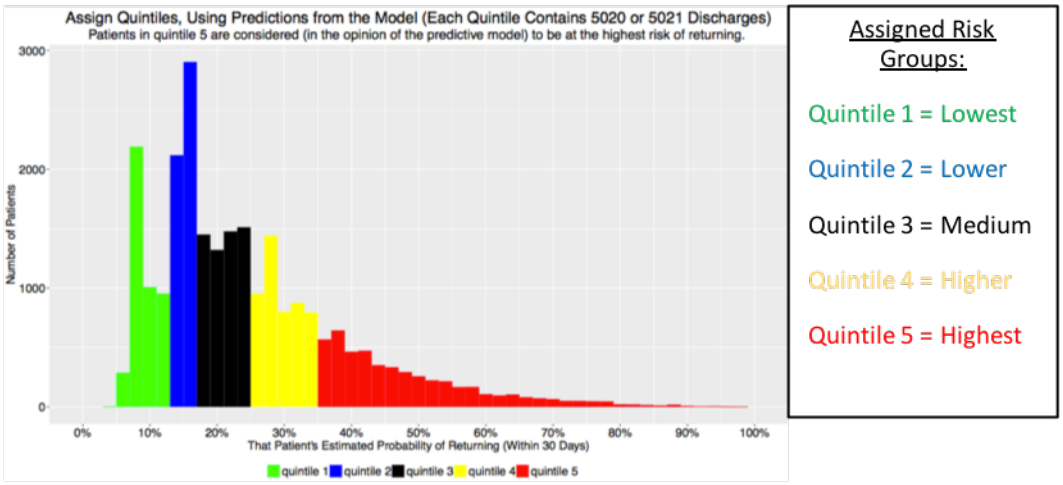
Figure 5C demonstrates the final cutoffs that were set in order to transform a calculated logitP into a risk score. These cutoffs were tested for accuracy within a 25,000 person test dataset of past UPMC Mercy patients that is pictured in Figure 5B, and quantified in 5C. From the graphical data in Figure 5B, it is evident that not all quintiles share an equal range; the 'lower' (quintile 2) has very narrow range, especially in comparison to the 'highest' (quintile 5) risk group, which ranges across more than half of the chart. As was expected, readmission rates amplified respectively across increasing risk quintiles (Supplemental Table 1).

A.

X_J	βJ if No Prior Hospital Visit	βJ if Prior Hospital Visit
Age	0.0038	-0.0052
Lpop	0.7259	0.4378
Female	-0.0770	-0.0310
Whether Source Is SNF	-0.5134	-0.8656
Whether Source Is Transfer	-0.2095	-0.2479
Whether Prior ED Visit	0.2255	0.0107
Whether Commercial Rollup	-0.0558	0.1509
Constant	-2.5992	-0.0052
Transformed Previous Hospital Visit	NA	0.5276

Lpop = Whether patient belongs to Low Socioeconomic status (SES)
SNF = Whether source of admission originated from a Skilled nursing facility
ED = Whether the Emergency Department is source of the admission
Transfer – Whether source of admission is a transfer from another hospital or unit
Commercial Rollup – Whether the patient is Medicare insured
Prior Visit – Patient was admitted within previous year
Transformed Previous Visit – Current admission directly stemmed from prior admission

B.



C.

Quintile	Number of Discharges	Minimum Prediction in that Quintile	Maximum Prediction in that Quintile	Average Return Rate (Out of Sample)
quintile 1	5,021	4.7%	13.6%	10.3%
quintile 2	5,021	13.6%	17.6%	15.4%
quintile 3	5,021	17.6%	24.7%	21.0%
quintile 4	5,021	24.7%	34.9%	29.7%
quintile 5	5,020	34.9%	97.5%	48.2%

Figure 5. Multiple Logistic Regression Model and Cutoffs Utilized to Make Readmission Predictions

4.2 SAMPLE SIZE CALCULATIONS

Because the testing dataset confirmed that the model was accurate under ideal conditions, the next step was to test the real-time accuracy of the model in a clinical setting. In order to approach this research question, sample size calculations were performed first at 80% power and 5% type I error. These were calculated for the primary outcome of interest, the one-sample proportion tests:

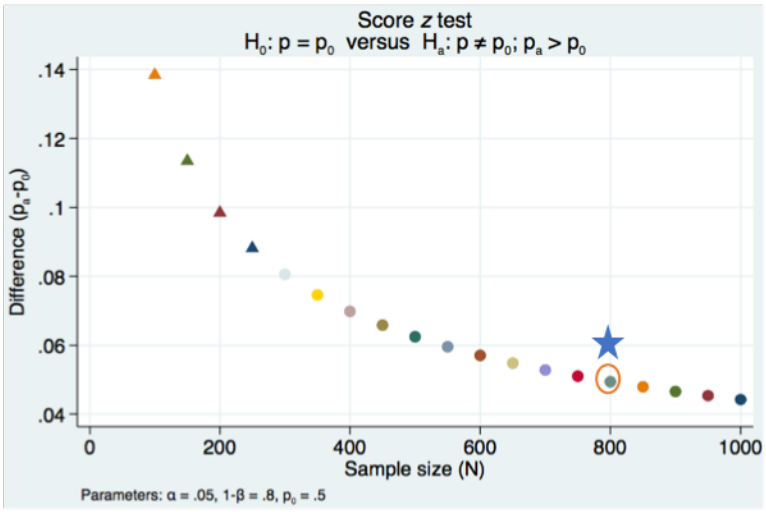
H_0: $p_{exp} = p_{obs}$

H_1: $p_{exp} \neq p_{obs}$

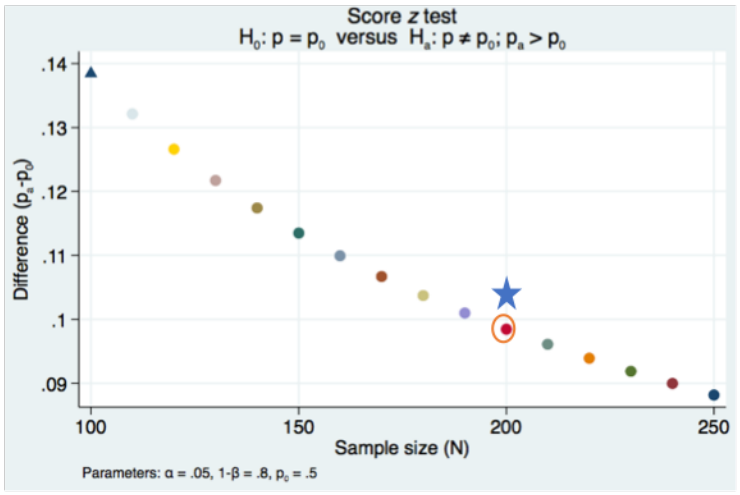
Figure 6A highlights an ideal sample size of recruiting 800 people per risk quintile, which would allow a ≥ 5.5% difference in expected and observed readmission rate to reject the null hypothesis. For this initial proof of concept study, we include 200 people per risk group, yielding a minimum rejectable difference between expected and observed of slightly under 10% (Figure 6B). While not as reliable of a range, it is worth noting that these projections are conservative. Figures 6A and 6B were calculated under the null hypothesis of $p_{exp} = 50\%$, which maximizes the number of people needed. In reality, no expected readmission rate was 50%, and most were around 15-20%, so these are conservative estimates to power the study. Figure 6C demonstrates this concept by putting different null hypotheses on the x-axis to display how the minimum detectable difference changes with n fixed at 200 people per group.

Figure 6 – Minimum Detectable Difference and Sample Size Calculations

A.



B.



C.

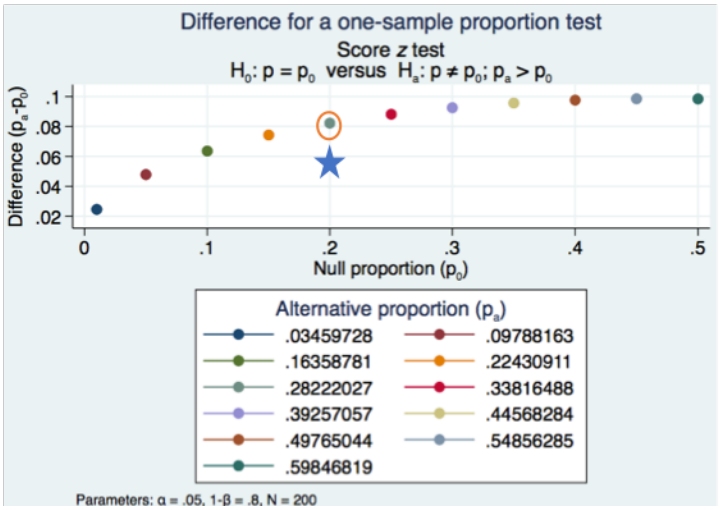


Figure 6. Minimum Detectable Difference and Sample Size Calculations

4.3 DEMOGRAPHIC ANALYSIS

Given that 200 people per risk group was the targeted sample size, a window of predictions that were made between April 7, 2018 and May 30, 2018 was chosen, consisting of 1,192 total patient predictions. Unfortunately, about 350 recruited individuals were censored prior to analysis due to duplicate predictions, missing patient information, and because some admissions were actually internal 'step-down' transfers without a true discharge. In the actual analyzed dataset, 858 patients remained, with only one risk group containing 200 patients. However, because the sample size calculations were overly conservative and that this was a proof of concept study, it was determined that the cohort was still of a sufficient sample size to continue onwards. After patient data was sorted from this time period, a brief demographic analysis was considered. It is important to note that no stratification can occur, because the logistic regression is already weighing these covariates. However, it still is very important to observe trends in covariates so that potential anomalies can be explained.

Table 2. Demographic Analysis By Risk Group Assignment

Risk Strata	Male (%)	Age (yrs.)	LOS (yrs.)
Lowest (n=125)	118 (94.4%)	68.688	10.232
Lower (n=189)	81 (42.9%)	60.873	8.063
Medium (n=200)	68 (33.85%)	52.12	5.303
Higher (n= 166)	82 (49.4%)	59.53	7.398
Highest (n= 178)	79 (44.4%)	58.393	7.713
Test Statistic	$\chi 2 = 125.71$	F = 18.15	F = 6.68
P-value	<0.0000	<0.0000	<0.0000
Significance	***	***	***

* Indicates a significant p-value ≤ 0.05
** Indicates a significant p-value ≤ 0.005
*** Indicates a significant p-value ≤0.0005
LOS – Length of Stay

Table 2 provides a cursory demographic analysis amongst the predicted readmission quintiles. Gender, age, and length of stay were all determined to be very significantly different across risk strata. The most glaring contributor to this observation is that 94.4% of the lowest risk strata consisted of male patients, which is highly unexpected. Supplemental Table 2 includes post-hoc ANOVA testing for LOS. Interestingly, all pairwise comparisons of the medium group had some level of significance. Supplemental Table 3 displays post-hoc ANOVA testing for Age, in which there are many significant pairwise comparisons. Simply looking at the trend however, the lowest risk strata has the highest mean age, and the medium risk strata has the lowest mean age, which again is unexpected.

4.4 χ^2 TESTING FOR INDEPENDENCE

The initial analysis intended to test if the assigned risk groups truly had distinct readmission rates from each other. Clinically, this test is incredibly important because if two risk groups experience overlapping mean readmission rates, then the predicted quintiles are not clinically useful for implementing interventions. First, a $\chi 2$ test of independence was conducted separately for 7-day and 30-day readmissions:

H_0: p1 = p2 = p3 = p4 = p5

H_1: At least one risk quintile has a different mean proportion of readmission

Table 3. $\chi 2$ Testing for Independence of Readmission Rates

	$\chi 2$ For All 5 Risk Groups	p-value
7-Day Readmission	18.396	0.001
30-Day Readmission	61.491	<0.000

As was hypothesized, Table 3 strongly favors the alternative hypothesis that for both 7-day and 30-day readmissions, at least one risk group experienced different rates of readmission than the rest.

Next, post-hoc analysis was conducted for 7-day readmissions. Crude p-values and Tukey adjustments for multiple comparisons are both shown in Table 4.

Table 4. 7-day Readmission $\chi 2$ Post-Hoc Analysis

	Crude p-value	Tukey Adjusted P-value	Significance
1 vs. 2	0.639	0.990	
1 vs. 3	0.128	0.547	
1 vs. 4	0.781	0.999	
1 vs. 5	0.025	0.164	*
2 vs. 3	0.025	0.164	*
2 vs. 4	0.414	0.925	
2 vs. 5	0.047	0.271	*
3 vs. 4	0.180	0.666	
3 vs. 5	<0.000	<0.000	**
4 vs. 5	0.006	0.050	**

* Indicates a significant crude p-value ≤ 0.05
** Indicates a significant adjusted p-value ≤ 0.05

Table 4 disputes hypotheses that all groups are distinct from one another, rather suggesting the opposite. When considering the adjustments for multiple comparisons, there are only 2/10 (20%) comparisons that are significantly different in mean 7-day readmission rates,

which both involved the 'highest' risk group. This suggests that the model is not an accurate predictor of 7-day readmission rates.

Following this analysis, 30-day readmission post-hoc $\chi 2$ analysis was considered. Table 5 demonstrates that the logistic regression model was much more accurate in predicting 30-day readmission, though still not all groups are distinct as was hypothesized. In Table 5, 5/10 (50%) head-to-head comparisons resulted in discrete mean rates of 30-day readmission. In examining which head-to-head comparisons are statistically significant, it is clear that the 'lowest' and 'highest' strata both overlap with other risk groups but remain distinct from each other as would be expected.

Table 5. 30-day Readmission $\chi 2$ Post-Hoc Analysis

	Crude p-value	Tukey Adjusted P-value	Significance
1 vs. 2	0.646	0.991	
1 vs. 3	0.178	0.661	
1 vs. 4	0.006	0.044	**
1 vs. 5	<0.000	<0.000	**
2 vs. 3	0.042	0.248	
2 vs. 4	0.010	0.072	*
2 vs. 5	<0.000	<0.000	**
3 vs. 4	<0.000	<0.000	**
3 vs. 5	<0.000	<0.000	**
4 vs. 5	0.009	0.071	

* Indicates a significant crude p-value ≤ 0.05
** Indicates a significant adjusted p-value ≤ 0.05

4.5 ONE-SAMPLE PROPORTION TESTS

Finally, one-sample proportion tests were conducted to assess how accurately the model's predictions followed patterns observed from the 25,000-person test dataset:

H0: $p_{Exp} = p_{Obs}$

H1: $p_{Exp} \neq p_{Obs}$

All expected mean readmission rates originate from analysis of the test dataset, which are chronicled in Supplemental Table 1.

First, observed 7-day readmission rates were compared to expected 7-day readmission rates. Table 6 demonstrates again that the logistic regression model is a poor predictor of 7-day readmission. Only the 'lowest' risk group did not show a statistically different rate of observed and expected readmission. Moreover, what is alarming about Table 6 is that observed readmission does not follow an increasing trend across higher strata, further highlighting that the 7-day risk groups do not have statistical or clinical significance. However, one positive aspect of Table 6 is the observed rate of the 'highest' risk group. Though not statistically similar to the expected mean rate of readmission, the 'highest' risk group still displayed a markedly higher mean rate of 7-day readmission than the other risk quintiles. From a clinical perspective, it is encouraging that the model can correctly predict the highest risk group for 7-day readmissions.

Table 6. 7-day Readmission One-Sample Hypothesis Test of Proportions

	N	Expected Readmission %	Observed Readmission %	P-value	Significance
Lowest	125	8.70%	5.60%	0.2188	–
Lower	189	13.90%	6.88%	0.0052	*
Medium	200	13.90%	1.49%	<0.0000	***
Higher	166	23.20%	4.82%	<0.0000	***
Highest	178	40.10%	11.80%	<0.0000	***

* Indicates a significant p-value ≤ 0.05
** Indicates a significant p-value ≤ 0.005
*** Indicates a significant p-value ≤0.0005

Next, 30-day expected and observed readmissions were compared to each other. Table 7 reinforces that the 30-day risk group predictions have much higher accuracy than the 7-day predictions. Statistically, the observed 'lowest', 'lower', and 'higher' risk strata are not significantly different from the expected rates of 30-day readmission. More encouragingly, there is high clinical relevance present here; observed 30-day readmissions increase across strata in an expected pattern with the exception of the 'medium' risk group. For example, the observed 'highest' risk group has a significantly different mean 30-day readmission from the expected value. However, clinically 36% is by far the highest observed readmission rate across all strata, and is much higher than the average rate of readmission. In this sense, the logistic regression model provides clinically useful predictions in regard to 30-day readmissions.

Table 7. 30-day Readmission One-Sample Hypothesis Test of Proportions

	N	Expected Readmission %	Observed Readmission %	P-value	Significance
Lowest	125	10.30%	12.80%	0.3578	–
Lower	189	15.40%	14.80%	0.8192	–
Medium	195	21.00%	6.97%	<0.0000	***
Higher	166	29.70%	25.30%	0.2147	–
Highest	178	48.20%	36.00%	0.0011	**

* Indicates a significant p-value ≤ 0.05
** Indicates a significant p-value ≤ 0.005
*** Indicates a significant p-value ≤0.0005

5.0 DISCUSSION

5.1 CONCLUSIONS

In a readmission prediction modeling validation study amongst patients at UPMC Mercy who were admitted between April 7 – May 3, 2018, we found mixed results with regards to the accuracy of the logistic regression readmission prediction tool. We hypothesized that this tool would successfully assign risk predictions such that increasing risk quintiles would experience increasing and distinct rates of 7-day and 30-day readmission. We also hypothesized that the risk quintiles in the retrospective cohort would experience similar mean rates of 7-day and 30-day readmission to the mean rates of readmission from the training dataset.

Sample size calculations determined that for a pilot study such as this, recruiting about 200 people per risk group would conservatively provide enough people to power the study at 80%, with a minimum detectable difference slightly below 10%. Patients who were admitted at UPMC Mercy Hospital between April 7, 2018 and May 30, 2018 were recruited to the study. Though seasonal variability in hospital visits could explain a small portion of the variation that exists within this study, it would be a source of nondifferential bias across the risk groups. Furthermore, it would not legitimately explain some of the extreme and unexpected observations such as the high male percentage in the 'lowest risk group' or the very low rates of readmissions within the 'medium' risk group.

Chi-square analyses of independence were conducted to test if the different risk quintiles experienced different mean rates of readmission from each other. When it was confirmed that the quintiles did have at least one risk group with different 7-day and 30-day readmissions from the others, pairwise post-hoc testing was considered. 7-day readmission post-hoc testing determined that there was a large amount of overlap in mean readmission rates amongst risk groups, with only 20% of comparisons showing a significantly different rate of readmission. When 30-day readmission underwent post-hoc chi-square testing, more groups were significantly distinct, though half of comparisons still overlapped with each other.

Finally, one-sample proportion tests were conducted for each individual stratum to compare expected rates of readmission from the training dataset and observed rates of readmission within this cohort. When 7-day readmissions were analyzed, only one risk quintile exhibited similar characteristics as the training dataset, further demonstrating that the model is not accurate when predicting 7-day readmissions. However, analysis of 30-day readmissions displayed mixed results. Statistically, only 60% of the quintiles followed readmission rates that were similar to the expected rates of readmission, which is still less than ideal. Nonetheless, 30-day readmission rates increased across risk groups as would be expected, with exception of the 'medium' risk group.

The multiple logistic regression model certainly did not perform to hypothesized expectations, and would likely benefit from another round of analyses that tweak the value of the beta coefficients and the cutoffs between quintiles. It is significant to note that when setting the cutoffs for the risk groups, the UPMC clinical analytics team concluded that there was high sensitivity and specificity for both the model's 7-day and 30-day readmission predictions. Before moving forward with the model, it is very important to understand why the model was accurate

with its 7-day readmission predictions in the training dataset, but not in our cohort. It is plausible that 7-day readmission is more variable than 30-day readmission because it involves a smaller follow-up time. Like with all statistics, as the sample size increases, the variation within the observations narrow. So, as the follow-up time increases, it is more likely that a study will detect more reliable mean rates of readmission.

The model in its current state clearly identifies the 'highest' risk group for both 7-day and 30-day readmissions. Because the intended purpose of the model is to target the highest risk patients for intervention, it may not be necessary to improve the accuracy of the model within lower risk groups. In its current iteration, these analyses support that the predictive model can effectively target the highest risk populations.

5.2 LIMITATIONS & POTENTIAL SOURCES OF BIASES

One potential source of bias is the discrepancies that may arise between the model and human data entry. Misclassification of outcome could affect the data here, because data entry may score readmissions differently than the model. For example, data entry could rate an internal transfer across UPMC as no readmission, whereas the model may score that as discharge and readmission on the same day. Another example could be different scoring of observational stays. Data entry was not checked by a panel, so it is also subject to human error.

Another limitation of the study could result from different hospitals treating the same patient. Studies have documented that between 20-40% of readmission occur at a different hospital from the index case[16, 20, 26, 27]. Additionally, if patients feel they received suboptimal care during an index admission, they will likely seek a different hospital upon readmision[27]. Hospitals

are only capable of tracking readmissions back to themselves, so there is likely a differential underestimate of readmission rates within this study[27]. However, because there are only two major healthcare systems in the greater Pittsburgh area, it is likely that it this underestimation is mild.

The demographic table provides a look at a few troubling sources of bias in our study. The medium risk group in our study consistently exhibited the lowest rate of 7-day and 30-day readmission. In the demographic analysis, the medium risk group also had the lowest male population percentage, shortest length of stay average, and youngest average age. It is worth further investigating why this occurred, and if it contributed to any introduction of bias.

One final source of bias could be an existing intervention effort at UPMC Mercy. There was a belief among some clinical and analytics professionals that people belonging to the lowest risk group did not need any special attention, and that those in the highest risk group were too unlikely to actually improve. With that methodology, a transitory care intervention was begun at UPMC Mercy that focused on patients identified as medium risk. Though further analysis needs to be conducted, a successful intervention effort potentially is what caused to the unexplained low trends in readmission rates in the medium risk group that is observed throughout the entirety of this study.

5.3 FUTURE DIRECTIONS

5.3.1 Potential Ways to Improve Model Accuracy

In order to improve the accuracy of the model, the first ideal step would be to continue recruiting more participants to the current study. The training dataset enrolled 25,000 patients in comparison to our 858 patients enrolled. There is a high likelihood that as we continue to add patients to the current trial, we would begin to normalize on the true mean readmission rates that are closer to the rates observed in the training dataset. Additionally, instead of relying on arbitrary cutpoint categorizations, it likely could be beneficial to conduct similar analyses that classify patients based on their exact continuous P(x) likelihoods for readmission.

The other potential way to improve the model would be to consider nontraditional risk factors in univariate and multivariate analyses. Both Kansangara et al. and Zhou et al. performed systematic reviews to characterize successful predictive models for readmission and found very limited success with very similar risk factors considered in each model, similar to what was utilized in our own model[18, 42, 50]. They suggest that overlooked aspects such as health literacy, family support, and availability of transportation can be combined with traditional measures such as LOS to generate a more accurate representation of predictive health[18].

5.3.2 Discharge Intervention Experimental Design

The next step in this study is to link specific transitory discharge interventions with predicted risk groups. This study can be completed in two different ways, which are both outlined in Figure 7. The most ideal scenario would be to conduct a randomized controlled trial (RCT), in

which patients within the same strata are matched for confounding risk factors, and then randomized to receive standard discharge protocols or the intervention discharge procedures. Following 30-days post-discharge, the 7-day and 30-day readmission rates can be calculated and analyzed. With this experimental design, the differences in readmissions rates across strata can be attributed to the intervention, allowing researchers to determine which readmission risk group intervention efforts benefit the most. If the resources are not available to conduct an RCT, a very similar analysis is detailed in Figure 4 utilizing a prospective cohort design.

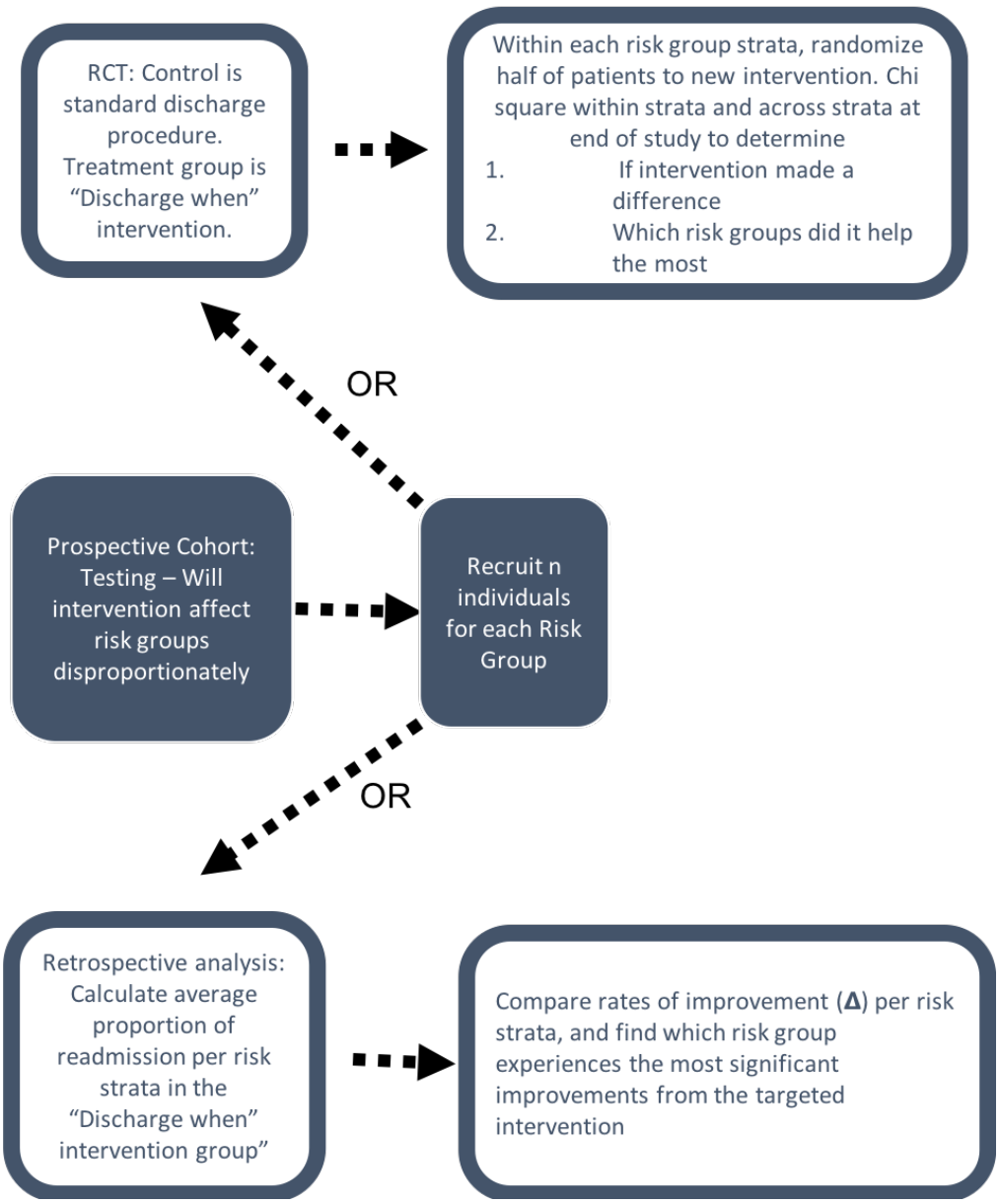


Figure 7. Future Discharge Intervention Experiment

6.0 PUBLIC HEALTH IMPACT

Readmissions are burdensome to patients, providers, and hospitals. Because hospital resources are limited, any model that can target high risk individuals is beneficial to public health outcomes. Frequent readmissions can be deleterious to a patient, causing a post-hospital syndrome in patients that results in heightened vulnerability for adverse outcomes, stemming from sleeplessness, anxiety, and stress associated with hospital admissions. Furthermore, the average cost of an admission is $13,100. But, the average cost of a readmission is $13,800 in addition to the cost of the index stay, which can further add stress and vulnerability to a patient.

Numerous interventions have demonstrated marked decreases in readmissions rates, but in most instances hospitals do not have the staffing resources to carry out discharge interventions systemically. Due to this, an accurate model that can target high risk patients would very likely improve patient outcomes and reduce readmission rates. Our prediction model specifically operates in real-time, which provides it a genuine opportunity to be utilized as a diagnostic tool. Since the intention of the prediction model is to aid decisions prior to patient discharge, a flagged high-risk patient can receive heightened patient education, medication reviews, planned transitory care, and communication with a healthcare professional.

APPENDIX: SUPPLEMENTAL TABLES

Supplementary Table 1. Readmission Rates from Test Dataset

Predicted Risk	7-day	30-day
Lowest	8.70%	10.30%
Lower	13.90%	15.40%
Medium	13.90%	21.00%
Higher	23.20%	29.70%
Highest	40.10%	48.20%

Supplementary Table 2. LOS Demographics ANOVA Post-Hoc Testing

	Crude p-value	Tukey Adjusted P-value	Significance
1 vs. 2	0.03	0.19	*
1 vs. 3	<0.000	<0.000	**
1 vs. 4	0.006	0.046	**
1 vs. 5	0.013	0.093	*
2 vs. 3	0.001	0.013	**
2 vs. 4	0.473	0.953	–
2 vs. 5	0.699	0.995	–
3 vs. 4	0.018	0.126	*
3 vs. 5	0.006	0.047	**
4 vs. 5	0.74	0.997	–

* Indicates a significant crude p-value ≤ 0.05
** Indicates a significant adjusted p-value ≤ 0.05

Supplementary Table 3. Age Demographics ANOVA Post-Hoc Testing

	Crude p-value	Tukey Adjusted P-value	Significance
1 vs. 2	<0.000	0.001	**
1 vs. 3	<0.000	<0.000	**
1 vs. 4	<0.000	<0.000	**
1 vs. 5	<0.000	<0.000	**
2 vs. 3	<0.000	<0.000	**
2 vs. 4	0.468	0.95	–
2 vs. 5	0.172	0.65	–
3 vs. 4	<0.000	0.001	**
3 vs. 5	<0.000	0.004	**
4 vs. 5	0.545	0.974	–

* Indicates a significant crude p-value ≤ 0.05
** Indicates a significant adjusted p-value ≤ 0.05

BIBLIOGRAPHY

1. Holloway, J.J. and J.W. Thomas, *Factors influencing readmission risk: implications for quality monitoring.* Health care financing review, 1989. **11**(2): p. 19-32.
2. Benbassat, J. and M. Taragin, *Hospital readmissions as a measure of quality of health care: Advantages and limitations.* Archives of Internal Medicine, 2000. **160**(8): p. 1074-1081.
3. Greenblatt, D.Y., et al., *Readmission after colectomy for cancer predicts one-year mortality.* Ann Surg, 2010. **251**(4): p. 659-69.
4. Jiang, H.J. and L.M. Wier, *All-Cause Hospital Readmissions among Non-Elderly Medicaid Patients, 2007: Statistical Brief #89*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).
5. Goldfield, N.I., et al., *Identifying Potentially Preventable Readmissions.* Centers For Medicare & Medicaid Services, 2008.
6. Kassin, M.T., et al., *Risk factors for 30-day hospital readmission among general surgery patients.* J Am Coll Surg, 2012. **215**(3): p. 322-30.
7. Horwitz, L.I., et al., *Development and Validation of an Algorithm to Identify Planned Readmissions From Claims Data.* Journal of hospital medicine, 2015. **10**(10): p. 670-677.
8. Hines, A.L., et al., *Conditions With the Largest Number of Adult Hospital Readmissions by Payer, 2011: Statistical Brief #172*. 2014: Agency for Healthcare Research and Quality (US), Rockville (MD).
9. Tahhan, G., et al., *Characterization of Planned and Unplanned 30-Day Readmissions Following Vascular Surgical Procedures.* Vascular and Endovascular Surgery, 2016. **51**(1): p. 17-22.
10. Mull, H.J., et al., *Association of postoperative readmissions with surgical quality using a delphi consensus process to identify relevant diagnosis codes.* JAMA Surgery, 2018. **153**(8): p. 728-737.
11. van Walraven, C., et al., *Proportion of hospital readmissions deemed avoidable: a systematic review.* Cmaj, 2011. **183**(7): p. E391-402.
12. Vu, T., L. Day, and C.F. Finch, *Linked versus unlinked hospital discharge data on hip fractures for estimating incidence and comorbidity profiles.* BMC medical research methodology, 2012. **12**: p. 113-113.
13. Burgess, J.F. and J.M. Hockenberry, *Can all cause readmission policy improve quality or lower expenditures? A historical perspective on current initiatives.* Health Economics, Policy and Law, 2014. **9**(2): p. 193-213.
14. Horwitz, L., et al., *Hospital-Wide (All-Condition) 30-Day Risk-Standardized Readmission Measure* Centers For Medicare & Medicaid Services, 2011.
15. CMS. *Hospital Readmissions Reduction Program (HRRP) Archives*. 2018.

16. Miller, M.E. and U. Senate, *Report to the Congress: Reforming the delivery system.* Washington, DC: Medicare Payment Advisory Commission, 2008. **2014**.
17. Fingar, K.R. and e. al., *A Comparison of All-Cause 7-Day and 30-Day Readmissions, 2014*, in *HCUP Statistical Brief #230*, M.S. Marguerite L. Barrett, and H. Joanna Jiang, Ph.D., Editor. 2017, Agency for Healthcare Research and Quality: Rockville, MD.
18. Kansagara, D., et al., *Risk prediction models for hospital readmission: A systematic review.* JAMA, 2011. **306**(15): p. 1688-1698.
19. CMS, *Hospital Readmissinos Reduction Program Frequently Asked Questions Fiscal Year 2019.* Centers For Medicare & Medicaid Services, 2019.
20. Ashton, C.M., et al., *The association between the quality of inpatient care and early readmission: a meta-analysis of the evidence.* Med Care, 1997. **35**(10): p. 1044-59.
21. Ashton, C.M., et al., *The association between the quality of inpatient care and early readmission.* Ann Intern Med, 1995. **122**(6): p. 415-21.
22. Miller, R.D.J., et al., *Readmissions Due to Hospital-Acquired Conditions (HACs): Multivariate Modeling and Undercoding Analyses.* Centers For Medicare & Medicaid Services, 2012.
23. Goodwin, A.J., et al., *Frequency, cost, and risk factors of readmissions among severe sepsis survivors.* Crit Care Med, 2015. **43**(4): p. 738-46.
24. CMS, *New Stratified Methodology Hospital-Level Impact File User Guide* 2017, Centers for Medicare & Medicaid Services: Hospital Readmissions Reduction Program.
25. Demiralp, B., F. He, and L. Koenig, *Further Evidence on the System-Wide Effects of the Hospital Readmissions Reduction Program.* Health Services Research, 2017. **53**(3): p. 1478-1497.
26. Jencks, S.F., M.V. Williams, and E.A. Coleman, *Rehospitalizations among Patients in the Medicare Fee-for-Service Program.* New England Journal of Medicine, 2009. **360**(14): p. 1418-1428.
27. Khan, A., et al., *Same-hospital readmission rates as a measure of pediatric quality of care.* JAMA Pediatrics, 2015. **169**(10): p. 905-912.
28. NQF, *All-Cause Admissions and Readmissions 2015-2017 Technical Report.* National Quality Forum, 2017.
29. Weiss, A.J., A. Elixhauser, and C. Steiner, *Readmissions to U.S. Hospitals by Procedure, 2010: Statistical Brief #154*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).
30. CMS, *ACO #8 - Risk Standardized All Condition Readmissions* Centers For Medicare & Medicaid Services, 2015.
31. Fingar, K. and R. Washington, *Trends in Hospital Readmissions for Four High-Volume Conditions, 2009–2013: Statistical Brief #196*. 2015: Agency for Healthcare Research and Quality (US), Rockville (MD).
32. Goldberg, E.M., et al., *An analysis of diagnoses that drive readmission: What can we learn from the hospitals in Southern New England with the highest and lowest readmission performance?* R I Med J (2013), 2017. **100**(8): p. 23-28.
33. Bambhroliya, A.B., et al., *Estimates and temporal trend for us nationwide 30-day hospital readmission among patients with ischemic and hemorrhagic stroke.* JAMA Network Open, 2018. **1**(4): p. e181190.
34. Magee, G., et al., *Impact of Clostridium difficile-associated diarrhea on acute care length of stay, hospital costs, and readmission: A multicenter retrospective*

study of inpatients, 2009-2011. American Journal of Infection Control, 2015. **43**(11): p. 1148-1153.
35. Birkmeyer, J.D., et al., *Hospital quality and the cost of inpatient surgery in the United States.* Ann Surg, 2012. **255**(1): p. 1-5.
36. Barrett, M.L., et al., *All-Cause Readmissions by Payer and Age, 2009–2013: Statistical Brief #199*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).
37. Song, X., et al., *Productivity loss and indirect costs associated with cardiovascular events and related clinical procedures.* BMC health services research, 2015. **15**: p. 245-245.
38. Mitchell, R.J. and P. Bates, *Measuring health-related productivity loss.* Population health management, 2011. **14**(2): p. 93-98.
39. Schoen, C. and et al., *Why not the best? Results from the national scorecard on US health system performance, 2008.* The Commonwealth Fund., 2006.
40. Borza, T., et al., *Association of the hospital readmissions reduction program with surgical readmissions.* JAMA Surgery, 2018. **153**(3): p. 243-250.
41. Zuckerman, R.B., et al., *Readmissions, Observation, and the Hospital Readmissions Reduction Program.* New England Journal of Medicine, 2016. **374**(16): p. 1543-1551.
42. Mendelson, A., et al., *The effects of pay-for-performance programs on health, health care use, and processes of care: A systematic review.* Annals of Internal Medicine, 2017. **166**(5): p. 341-353.
43. Desai, N.R., et al., *Association between hospital penalty status under the hospital readmission reduction program and readmission rates for target and nontarget conditions.* JAMA, 2016. **316**(24): p. 2647-2656.
44. Hansen, L.O., et al., *Interventions to reduce 30-day rehospitalization: A systematic review.* Annals of Internal Medicine, 2011. **155**(8): p. 520-528.
45. Baker, C.D., et al., *A Standardized Discharge Process Decreases Length of Stay for Ventilator-Dependent Children.* Pediatrics, 2016. **137**(4): p. e20150637.
46. Elkassabany, N.M., et al., *Hospital Characteristics, Inpatient Processes of Care, and Readmissions of Older Adults with Hip Fractures.* Journal of the American Geriatrics Society, 2016. **64**(8): p. 1656-1661.
47. Chang, A.M. and K.L. Rising, *Cardiovascular Admissions, Readmissions, and Transitions of Care.* Current emergency and hospital medicine reports, 2014. **2**(1): p. 45-51.
48. Stelfox, H.T., et al., *Assessment of the safety of discharging select patients directly home from the intensive care unit: A multicenter population-based cohort study.* JAMA Internal Medicine, 2018.
49. Costantino, M.E., et al., *The influence of a postdischarge intervention on reducing hospital readmissions in a Medicare population.* Population health management, 2013. **16**(5): p. 310-316.
50. Zhou, H., et al., *Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review.* BMJ open, 2016. **6**(6): p. e011060-e011060.
51. Burke, R.E. and E.A. Coleman, *Interventions to decrease hospital readmissions: keys for cost-effectiveness.* JAMA Intern Med, 2013. **173**(8): p. 695-8.
52. Steiner, C., M. Barrett, and K. Hunter, *Hospital Readmissions and Multiple Emergency Department Visits, in Selected States, 2006–2007: Statistical Brief #90*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).

53. Barnett, M.L., J. Hsu, and J.M. McWilliams, *Patient Characteristics and Differences in Hospital Readmission Rates.* JAMA internal medicine, 2015. **175**(11): p. 1803-1812.
54. Wier, L.M., et al., *All-Cause Readmissions by Payer and Age, 2008: Statistical Brief #115*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).
55. Bashir, B., et al., *Evaluation of prediction strategy and care coordination for COPD readmissions.* Hospital Practice, 2016. **44**(3): p. 123-128.
56. Yeh Robert, W., et al., *Sources of Hospital Variation in Short-Term Readmission Rates After Percutaneous Coronary Intervention.* Circulation: Cardiovascular Interventions, 2012. **5**(2): p. 227-236.
57. Siracuse, B.L. and R.S. Chamberlain, *A preoperative scale for determining surgical readmission risk after total hip replacement.* JAMA Surgery, 2016. **151**(8): p. 701-709.
58. Elixhauser, A. and C. Steiner, *Readmissions to U.S. Hospitals by Diagnosis, 2010: Statistical Brief #153*. 2006: Agency for Healthcare Research and Quality (US), Rockville (MD).
59. Campione, J.R., S.A. Smith, and R.E. Mardon, *Hospital-Level Factors Related to 30-Day Readmission Rates.* American Journal of Medical Quality, 2015. **32**(1): p. 48-57.