USE OF ADMINISTRATIVE CLAIMS DATA TO DESIGN AND EMULATE A CLINICAL TRIAL IN ACUTE STROKE PATIENTS COMPARING REHABILITATION AT INPATIENT REHABILITATION FACILITIES TO SKILLED NURSING FACILITIES

By

Kent P. Simmonds

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Epidemiology-Doctor of Philosophy

2020

ABSTRACT

USE OF ADMINISTRATIVE CLAIMS DATA TO DESIGN AND EMULATE A CLINICAL TRIAL IN ACUTE STROKE PATIENTS COMPARING REHABILITATION AT INPATIENT REHABILITATION FACILITIES TO SKILLED NURSING FACILITIES

By

Kent P. Simmonds

Stroke affects nearly 800,000 people every year in the United States and is a leading cause of adult disability. After hospitalization half of stroke patients continue to require medical and rehabilitation services provided at inpatient rehabilitation facilities (IRFs) or skilled nursing facilities (SNFs). In general, IRFs provide time-intensive therapy for two to three weeks, while SNFs provide moderately intensive therapy for four- to five-weeks. There is substantial variation in the utilization of these alternative rehabilitation settings, but their relative comparative effectiveness remains uncertain. A randomized controlled trial (RCT) would provide an unbiased comparative effectiveness estimate, but the design of such a trial is complicated by several practical and ethical issues. The overarching purpose of this dissertation was to use Medicare claims data to inform the design and to emulate such a trial.

<u>In the first aim</u>, we sought to identify patient and hospital level factors that were associated with IRF or SNF discharge and characterize the heterogeneity of hospital effects that influenced discharge to an IRF (vs. SNF). From a retrospective cohort of 145,894 stroke patients, we used multi-level multivariable models to identify several patient- and hospital- level factors that were independently associated with discharge setting. We also showed that hospitals contributed around a third of the variation in IRF (vs. SNF) discharge, but there was substantial variation in the effect that specific hospitals had on influencing IRF discharge.

The second aim, was to identify a target trial population that optimized the explanatory-pragmatic balance of a subsequent RCT. To identify this population, we profiled hospitals based on their propensity to discharge stroke patients to IRFs (vs. SNFs) and inferred IRF and SNF referral networks for each hospital. The final target trial population included 44,950 patients (30.8% of the starting sample) who were treated at 441 hospitals (14.5%) and subsequently discharged to 745 IRFs (64.8%) and 5,974 SNFs (48.2%).

The third aim was to emulate three alternate RCTs that compared patient outcomes at IRFs vs. SNFs. Trial #1 used the target trial population identified in Aim 2, while trials #2 and 3 excluded increasingly infrequently used IRFs and SNFs. Comparative effectiveness was estimated using a matched propensity score analysis. Overall, on a relative basis, patients treated at IRFs were between 18-35% more likely to be successfully discharged home (i.e., alive and at home for >30 days) and were between 11-15% less likely to die within one year of acute care discharge. The variation in the effect size estimates across the trials was driven by poorer outcomes among patients treated at infrequently used SNFs. Finally, we identified that a moderate sized unmeasured confounder would nullify the observed differences.

In conclusion, we identified that referring hospitals are a major driver of IRF or SNF use, and that patients treated at IRFs had better outcomes (relative to SNF patients). However, our results were limited by the inability to adjust for potentially important unmeasured confounders. A pragmatic RCT would eliminate such biases and provide a more valid comparative effectiveness estimate of these two alternative rehabilitation settings.

ACKNOWLEDGEMENTS

The completion of this dissertation was an incredible journey that could not have been completed without the guidance, support, and encouragement from numerous people in my life. First, I would like to thank my advisor, Dr. Mathew Reeves whose expectations pushed me to develop much more critical, rigorous thinking while also ensuring that I communicated with clarity and precision. In addition, I would like to thank the other members of my committee: Drs. Michael Andary, James Burke, Allan Kozlowski, and Zhehui Luo. Their combined input and experience were vital to guide this project in a more productive and rigorous direction. I would also like to thank the DO-Ph.D. program at Michigan State, notably Dr. Justin McCormick for initially believing in my potential and Dr. Brian Schutte for the continued institutional support. In addition, I would like to thank the faculty, staff, and students from the Epidemiology and Biostatistics program for their knowledge, support and camaraderie.

I would also like to thank my parents-especially my mom who has served as an academic role model throughout my life and inspired me to pursue a career in research. In addition, I would also like to thank Natasha for her eternal love, encouragement, and support and Chad for his encouragement and optimism. Finally, I would like to thank Dr. Brian Degenhardt for his belief and encouragement that I had the ability to pursue a career as a physician scientist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

IRF(s)	inpatient rehabilitation facility(ies)
SNF(s)	skilled nursing facility (ies)
RCT(s)	randomized controlled trial(s)
LOS	length of stay
PAC	post-acute care
CMS	Centers of Medicare and Medicaid Services
OR	odds ratio
aOR	adjusted odds ratio
IPW(s)	inverse probability weight(s)
PS	propensity score
CI	confidence interval
FIM	Functional Impairment Measure (FIM)
MO	health maintenance organization
IV(s)	instrumental variable(s)
ICF	International Classification of Functioning, Disability, and Health
MRI(s)	magnetic resonance image(s)
t_0	time zero
ICD-9	International Classification of Diseases, Ninth Revision

CHAPTER 1: BACKGROUND AND OBJECTIVES

Every year in the United States, approximately 800,000 people experience a stroke.[1] Stroke is the 5th leading cause of death (140,000 annual deaths) and is the leading cause of adult disability.[1] Acute hospital care for stroke patients is often short (i.e. the average length of stay (LOS) is 4 days), focuses on stabilizing the patient, and survivorship is high (~95%).[2] However, stroke survivors often face numerous short and long term health issues which can be either a direct result of their stroke (e.g., physical disability, cognitive impairment, seizures, pain) or from complications arising from their stroke (e.g., urinary tract infections, decubitus ulcers, depression).[1] These health issues are both deadly (30 day mortality is 15%) and disabling.[1,3] To manage these health issues, improve functions of activity and mobility, and support the transition back to the community most stroke patients require some form of post-acute care (PAC). The organization of PAC is highly heterogenous with numerous settings offering different types and intensities of treatments.[4] The most common PAC settings include inpatient rehabilitation facilities (IRFs), skilled nursing facilities (SNFs), home health care, and outpatient rehabilitation.[5,6]

Around half of stroke patients will be discharged from acute care to receive IRF (~25%) or SNF (~25%) care.[5,6] Rehabilitation at IRF and SNFs is often compared because both are inpatient PAC settings that primarily focus on rehabilitation, roughly equal numbers of patients receive each type of rehabilitation care, and some physicians consider each type of care as interchangeable.[7,8] However, there are several significant regulatory differences in the minimum levels of therapy, facility structure, and the extent of clinical oversight provided at each setting. IRFs provide hospital level care with daily oversight of physicians and registered nurses are available 24 hours a day.[4] Patients receive time-intensive (minimum of 3 hours per day) rehabilitation therapy delivered over a LOS of two to three weeks.[9,10] SNFs are often freestanding nursing homes that provide a broad range of clinical oversight and therapy intensities. Although physicians oversee a treatment plan, they rarely have daily contact with patients, and nurse availability ranges from 8 to 24 hours a day.[4] Patients who receive SNF care often receive a wider range of therapy intensities (i.e. 45-720 minutes per week) with a typical LOS of three to five weeks.[10] However, over the course of a typical IRF or SNF stay the total therapy time is similar for the two populations.[6]

Clinically, based off the 2016 Stroke Rehabilitation guidelines, IRF care is indicated for high acuity patients who are able to tolerate high intensity therapy for at least 3 hours a day for 5 days per week. In addition, IRFs are indicated for patients wo have an expectation of significant improvements of mobility and self-care activities, and are anticipated to be discharged back to the community.[4] Other patient indications for IRF care include regular physician contact for medical comorbidity management, and complex rehabilitation needs (e.g. orthotics, spasticity, acute illness).[4] In contrast, SNFs are indicated for lower acuity patients with less complex health care needs and are only expected to make partial recovery.[4] Other indications for SNF care include patients who require nurses to manage and prevent further health deterioration for pre-existing conditions such as decubitus ulcers, bowel, bladder impairment, or are at risk for nutritional deficiencies.[4]

Identifying which patients are best suited for IRF or SNF care is challenging for several reasons.[8] First, stroke recovery is highly heterogenous and many personal (e.g. age, sex, insurance, social support), clinical (e.g. stroke severity, comorbidities, physical and cognitive function), and environmental factors (e.g. home environment, access to follow-up medical care) can affect both recovery trajectories and the odds of discharge back to the community.[11,12]

Second, even if clinicians could reliably predict a patient's recovery prognosis, there remains an insufficient understanding of the relative effectiveness of IRF versus SNF care and whether there is heterogeneity of treatment effect. That is, it is unclear if the relative treatment effect size for IRF vs. SNF care is constant across a baseline spectrum of function- specifically function related to mobility, self-care, and cognition.[13,14,15]

Based on analyses of large national databases, previous studies have identified several significant sociodemographic, clinical, and environmental level differences in the characteristics of patients who were discharged to an IRF compared to patients discharged to a SNF.[15–17] In general, patients who received care at an IRF tended to be younger, male, had fewer comorbidities, and had lower health care utilization prior to their stroke (e.g., fewer hospitalizations). In addition, IRF patients tended to have less severe strokes and received less health care utilization (e.g., shorter LOS). Patients who attend IRFs were also more likely to reside in urban settings.[15–17] However, despite these differences, overall there is a sizable overlap in patient level characteristics between the two populations.[15–17]

A multitude of factors (many of which are independent of a patients clinical need) drive this sizable overlap between the IRF and SNF populations.[18] First, patients and their families are frequently consulted to identify a specific rehabilitation facility which would best meet their needs. Qualitative interviews have shown that factors such as financial resources, social support, motivation, are all important considerations for patients and their families.[19,20] Another patient centered factor includes selecting a facility close to a patients home. However, the geographic distribution of IRF and SNFs are not equivalent as there are ~15 times more SNFs compared to IRFs and SNFs tend to be much smaller and are diffusely spread across the United States.[21] In contrast, IRFs are often much larger and are clustered in urban areas.[15,22] Other significant factors that influence referral to specific facilities may include a patient's health insurance, hospital preferred referral networks, and bed availability at IRFs and SNFs.[9] Finally, several studies have identified that the acute care hospital which treated the acute stroke, has a very large effect on influencing discharge to an IRF or SNF.[15–17] Specific drivers of this effect remain poorly understood, but hospital referral networks, the hospital clinical culture, patient case-mix, and geographic availability of IRF or SNF bed availability likely all contribute to the variation in IRF and SNF use.[15–17,23]

At a time when healthcare expenditure has become a national priority, PAC use has garnered increased attention for several reasons. First, shorter acute hospital stays have shifted more patients to use PAC, and this has increased total PAC costs.[24] In 2001, the Centers for Medicare and Medicaid Services (CMS) spent $29.3 billion on PAC, but in 2017 total PAC costs increased to $58.5 billion.[21] This issue is central for IRF vs. SNF care for stroke rehabilitation, because stroke patients use of IRF and SNF care is fundamental to the variation in PAC spending because stroke patients are the second highest users of PAC and IRF care costs are approximately double SNF care (i.e. $19,149 versus $10,482) for stroke patients.[10] Second, there is significant regional variation in PAC spending, for CMS patient specifically, PAC was responsible for ~70% of the total variation in spending despite only accounting for ~30% of the total costs. [25,26] For IRF and SNF care for stroke patients, several studies have also found very large state-to-state variation in IRF (4-30%) and SNF (14 to 40%) use for stroke patients.[18,15]

COMPARATIVE EFFECTIVENESS OF IRF VERSUS SNF CARE

A recent systematic review of comparative effectiveness studies found that, in general stroke patients discharged to a IRF had better outcomes relative to those discharged to a SNF.[27] Noteworthy, this review was based on only seven studies all of which were observational in nature and used a variety of different patient outcomes including functional gain, community discharge, hospital readmissions, and/or mortality. The majority of the seven included studies used large administrative databases and used multivariable models to adjust for a variety of potential confounding factors.

Table 1.1 provides a summary of the details of seven studies that were included in the systematic review, and is organized according to the type of outcomes assessed (i.e., physical functioning, community discharge, hospital readmissions, and all-cause mortality). For community discharge, it can be seen that patients who received care at an IRF had around twice the odds of being discharge home.[10] However, the exact effect depended on the specific study population and statistical analysis which was performed. Three out of the four studies that assessed physical functioning, identified that IRF patients had improved function relative to SNF patients. Only one study compared hospital readmissions and found that IRF patients (vs. SNF patients) were marginally less likely to be readmitted to a hospital or the emergency department. Finally, all four studies that measured all-cause mortality found that overall IRF patients were less likely to die compared to SNF patients, but the largest differences in mortality were observed within the first 14 days after acute care discharge. However, for all outcomes the exact effect comparative effectiveness estimate depended on the specific study population and statistical analysis which was performed.

Table 1.1 Differences in health outcomes for stroke rehabilitation at IRFs or SNFs among acute stroke patients-table adapted from Alcusky et al. 2017

Author (Year)	**Data (Sample size)**	**Analytic Approach**	**Comparison**	**Crude Percentages or Means**	**Measure(s)**	**Summary of findings**
				Community Discharge		
Deutsch et al. (2006)	Uniform Data System for Medical Rehabilitation and Medicare Provider Analysis and Review (n=58,724)	Multivariable logistic model	IRF vs. SNF (Reference)	Stratified by admission disability level: Minimal motor: **IRF:** 98.6%; **SNF:** 98.6%	aOR 95% CI	Community discharges in IRF more common than in SNF for these patients:
				Mild motor/mild cognitive: **IRF:** 96.7%; **SNF:** 91.7% Minimal motor/significant cognitive: **IRF:** 90.6%; **SNF:** 88.3%		Mild motor disabilities and cognitive ratings: aOR: 2.19; 95% CI: 1.52–3.14
				Moderate motor: **IRF:** 92.3%; **SNF:** 84.2%		Moderate motor disabilities: aOR: 1.98; 95% CI: 1.49–2.61
				Significant motor: **IRF:** 85.8%; **SNF:** 79.3%		Significant motor disabilities: aOR: 1.26; 95% CI: 1.01–1.57
				Severe motor--patients ≥ 82 years: **IRF:** 54.6%; **SNF:** 49.4% patients < 82 years: **IRF:** 66.4%; **SNF:** 52.0%		Severe motor disabilities, patients <82 years: aOR: 1.43; 95% CI: 1.25–1.64

Table 1.1 (cont'd)

Hoenig et al. (2001)	Veterans' Health Administration databases (n=5,168)	Multivariable logistic model	Rehabilitation unit and geriatric unit vs. SNF (Reference)	**Rehabilitation unit:** 75.0% **Geriatric unit:** 71.8% **SNF**: 66.6%	aOR 95% CI	Relative to those in SNFs, patients in rehabilitation units (aOR: 1.91; 95% CI 1.47–2.50) and geriatric units (aOR:1.43; 95% CI 1.03–1.97) had increased odds of being discharged home.
Bettger et al. 2019	Medicare and Get with the Guidelines Cohort study data (n=162,423)	Propensity Scores Instrumental variable analysis	IRF vs. SNF	Total amount of home-time 90 days **IRF:** 51.8+/- 31.2 **SNF:** 32.5 +/- 30.7 365 days: **IRF:** 271.2 +/- 112.5 **SNF:** 195.5 +/- 138.5	Adjusted hazard ratios 95% CI to measure home-time	Propensity Score 90 day: 1.4 95% CI: 1.3-1.4 365 day: 1.2 95% CI: 1.2-1.2 IV (% IRF) 90 day: 1.3 95% CI: 1.20-1.3 365 day: 0.9 95% CI: 0.9-1.0 IV (Differential distance) 90 day: 1.18 95% CI: 1.0-1.4 365 day: 1.0 95% CI: 0.9-1.1

Table 1.1 (cont'd)

				Physical Functioning		
Chen et al. (2002)	Uniform Data System for Medical Rehabilitation (n=349)	Multiple linear model	IRF and Acute hospital vs. SNF (Reference)	Average Rasch-transformed Mobility Gain (range 0–100): 17	Standardized β Coefficient	Patients in SNFs made larger gains in mobility than patients in IRF (−0.20; p<0.05) or patients in acute hospitals (−0.16; p<0.05).
Deutsch et al. (2006)	Uniform Data System for Medical Rehabilitation and Medicare Provider Analysis and Review (n=58,724)	Multiple linear model	IRF vs. SNF (Reference)	Discharge FIM motor rating stratified by disability level: Minimal motor: **IRF:** 86.6; **SNF:** 85.0 Mild motor/mild cognitive: **IRF:** 79.2; **SNF:** 78.3 Minimal motor/significant cognitive: **IRF:** 77.5; **SNF:** 77.5 Moderate motor: **IRF:** 73.1; **SNF**: 71.1	Adjusted β coefficient representing the mean FIM difference (IRF-SNF) 95% CI	Clinically relevant functional gains (≥2 FIM units) in IRF more common than in SNF for these patients:
				Significant motor: **IRF:** 67.1; **SNF:** 64.9		Significant motor disabilities: adjusted β: 2.40; 95% CI: 1.19–2.66
				Severe motor-- patients ≥ 82 years: **IRF:** 46.1; **SNF:** 40.1 patients < 82 years: **IRF:** 49.8; **SNF:** 41.8		Severe motor disabilities—patients ≥82 years: adjusted β: 2.39; 95% CI: 1.45–3.32 patients <82 years: adjusted β: 4.24; 95% CI: 3.45–5.03

Table 1.1 (cont'd)

Physical Functioning						
Kane et al. (2000)	Medicare automated data retrieval system; patient survey, and medical records (n=202)	Multiple linear model Instrumental variable analysis	IRF vs. SNF (Reference)	Average percentage change in the activities of daily living score at six weeks, 6 months, and 12 months. Crude average change values were not provided. **IRF:** 6 weeks: 23.2% improved 6 months: 13.9% improved 12 months: 7.8% improved **SNF:** 6 weeks: 0.7% improved 6 months: −5.9% worsened 12 months: −6.7% worsened	Adjusted mean functional dependency scores Predicted gain in functional improvement in optimal post-acute care setting	IRF patients regained more activities of daily living at six weeks. Despite some rebound loss of activities of daily living between 6 and 12 months, IRF patients fared better than SNF patients SNF and IRF settings differed most at 6 weeks (**IRF:** 3.1%, **SNF:** 16.9%) and were similar at 6 months (**IRF:** 15.5%, **SNF:** 18.3%) and 12 months (**IRF:**15.9%, **SNF:** 16.2%)
Chan et al. (2013)	Kaiser Permanente Health Care System Northern California (222)	Multiple linear model	SNF vs. IRF (reference	AMPAC score at 6 months: **IRF:** 52 **SNF**: 43	Adjusted β coefficient representing the mean AM-PAC difference (SNF-IRF) 95% CI	Adjusting for hospital readmission and quantity of therapy (adjusted β: −10.1; 95% CI: −15.0 to −5.2), and adjusting for readmission and quantity of therapy (adjusted β: −6.1; 95% CI: −11.2 to −1.0).

Table 1.1 (cont'd)

			Hospital Readmission			
Kind et al. (2010)	Medicare Provider and Analysis Review; Provider of Services (31,283)	Unspecified statistical model with robust variance estimates to account for clustering of patients within hospitals	IRF and SNF	Crude estimates not available by site of care.	Predicted probability of readmission (hospital or emergency department) 95% CI	Predicted probabilities of readmission less for IRF than SNF in each racial/ethnic group. Blacks: **IRF:** 20%; 95% CI: 17.9–22.7 **SNF:** 26%; 95% CI: 24.2–28.6 Hispanics: **IRF:** 18%; 95% CI: 13.1–22.9 **SNF:** 28%; 95% CI: 24.0–32.6 Whites: **IRF:** 18%; 95% CI: 17.3–19.1 **SNF:** 21%; 95% CI: 20.3–21.9
			All-cause Mortality			
Buntin et al. (2010)	Medicare Provider and Analysis Review; Minimum Data Set (n=156,750)	Generalized estimating equations (binary logit) Instrumental variable analysis	IRF vs. SNF (Reference)	Mortality within 120 days **IRF:** 6.2% **SNF:** 14.7%	Absolute difference in 120-day mortality 95% CI	Use of IRF reduced mortality by 2.6 percentage points compared to SNFs. adjusted β: −2.58; 95% CI: 0.96–4.16

Table 1.1 (cont'd)

				All-cause Mortality		
Kind et al. (2010)	Medicare Provider and Analysis Review; Provider of Services (31,283)	Unspecified statistical model with robust variance estimates to account for clustering of patients within hospitals	IRF and SNF	Crude estimates of 30-day mortality not available by site of care.	Predicted probability of 30-day mortality among those with no readmissions 95% CI	Predicted probability of death in IRF settings lower than SNF settings in each racial/ethnic group. Blacks: **IRF:** 2%; 95% CI: 1.6–3.3 **SNF:** 5%; 95% CI: 4.2–6.1 Hispanics: **IRF:** 1%; 95% CI: 0–1.5 **SNF:** 5%; 95% CI: 3.2–6.3 Whites: **IRF:** 2%; 95% CI: 1.9–2.5 **SNF:** 8%; 95% CI: 7.2–8.2
Wang et al. (2011)	Kaiser Permanente California Health System Claims (n=17,348)	Cox proportional hazards multivariable model	IRF vs. SNF (Reference)	Stratified by the highest level of post-acute care within 14 and 61 days: *Post-acute (14 days)*: **IRF:** 4.4% **SNF:** 21.4% *Post-acute (61 days)*: **IRF:** 4.3% **SNF:** 16.2%	Adjusted hazard rate ratio 95% CI	Patients in IRF settings died at a rate less than half that of those in SNF settings. *Post-acute (14 days)*: Adjusted hazard ratio: 0.33 95% CI 0.24–0.45 *Post-acute (61 days)*: Adjusted hazard ratio: 0.42 95% CI 0.33–0.53

Table 1.1 (cont'd)

All-cause Mortality						
Bettger et al. 2019	Medicare and Get with the Guidelines Cohort study data (n=162,423)	Propensity Scores Instrumental variable analysis	IRF vs. SNF	14 days **IRF:** 1.1% **SNF:** 6.38% 90 days: **IRF:** 7.2% **SNF:** 21.1% 365 days: **IRF:** 17.9% **SNF:** 38.6%	Adjusted hazard ratios 95% CI	Propensity Score 14 day: 0.28 95% CI: 0.24-0.33 90 day: 0.52 95% CI: 0.49-0.55 365 day: 0.65 95% CI: 0.62-0.68 IV (% IRF) 14 day: 0.55 95% CI: 0.44-0.69 90 day: 0.77 95% CI: 0.70-0.85 365 day: 0.92 95% CI: 0.86-0.98 IV (Differential distance) 14 day: 0.31 95% CI: 0.17-0.57 90 day: 0.74 95% CI: 0.58-0.96 365 day: 0.89 95% CI: 0.75-1.05

Abbreviations: skilled nursing facility (SNF), inpatient rehabilitation facility (IRF), adjusted odds ratio (aOR), confidence interval (CI), Functional Impairment Measure ™ Instrument (FIM), health maintenance organization (HMO), Instrumental variables (IV)

Table adapted from Alcusky et al. 2017

% IRF: Based on percent of patients a hospital discharged to an IRF (vs. SNF), Differential Distance: Distance to the closest IRF – closest SNF from a patient's home (estimated as the midpoint of a patient's zip code)

Due to their observational designs, several methodological limitations should be considered when interpreting the results of these studies. First, there are many complex selection forces that act on the patient, hospital, and environmental levels which may guide patients with more favorable recovery prognoses to receive rehabilitation at IRFs (vs. SNFs). All of the studies attempted to control for these baseline differences via statistical adjustment. However, it is unclear how precise many important factors are able to capture complex issues such as medical acuity, and pre-and post-stroke function.[8,28] In addition, important unmeasured confounders (e.g., social support, patient motivation, provider biases towards appropriate care) were not captured.[8,28] Thus, from these studies it remains unclear how much of the observed effect is real and how much of the effect is due to residual confounding. A 2017 Medicare Payment Advisory Commission Report to the Congress summed up the limitations of observational comparative effectiveness studies as, "For any given patient the need for PAC is not clear, and there is limited evidence on which setting would be best and what mix of services would achieve the best outcomes."[13]

In addition to the methodological limitations of these studies there are several clinical reasons why uncertainty towards the relative effect of IRF versus SNF care exists. First, despite obvious differences in the type, intensity and duration of therapy at the two settings over the course of a typical stay (i.e. median 15 days for IRF patients and 35 days for SNF patients), the total amount of therapy received is quite similar.[6] Second, many studies focus on physical and activity level function as primary outcomes because the goals of IRF and SNF care are often aimed at restoring mobility, self-care, cognitive, and communication level functions to ultimately promote community living.[4] However, function is a complex construct and under the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) model, functional improvements could be driven by biological and/or clinical processes (i.e. tissue restoration, and neural reorganization), learning dependent processes (i.e. obtaining new skills to meet environmental demands), social factors (e.g., additional informal support from family, friends, or hired medical assistants), and/or environmental factors (e.g., disability transport, wheelchair ramp availability, etc.).[11,29] However, measuring the effect that differences in the timing, intensity, duration, type, and frequency of therapy delivered at IRF vs. SNFs has on each of these processes is difficult and the results are unclear.[8,11,30] Third, although theoretical benefits of IRF care over SNF care (i.e. close clinical monitoring, multidisciplinary rehabilitation teams) are clear, a 2018 report by the office of the inspector general found that the majority of IRF patients did not meet eligibility requirements for highly specialized rehabilitation care in a hospital setting.[31] In fact, many of these patients received rehabilitation services (e.g., general exercise, therapies targeted to improve ambulation or sitting tolerance) that could have been delivered in an outpatient setting.[31] However, these results were not specific for stroke patients who are among rehabilitation populations with some of the most substantial rehabilitation needs.

However, several clinical reasons may account for the observed improved outcomes for IRF patients. First, IRF patients have closer clinical monitoring due the 24/7 availability of nurses and daily physician access that could enable earlier treatment of complications. Second, although overall IRF and SNF patients receive a similar total amount of therapy time, IRF patients receive higher intensity rehabilitation for a shorter duration of time.[32,33] Higher intensity rehabilitation therapy received early in the recovery period may increase both the rate of return and maximum level of physiologic and activity function a patient achieves.[34–36] Third, patients treated at IRFs receive care from a multi-disciplinary team of highly trained rehabilitation professionals which may be better able to identify and manage a patient’s rehabilitation needs.

Finally, many IRFs are physically embedded within hospitals which provides access to hospital level medical equipment (e.g., magnetic resonance images (MRIs)) and other specialists (e.g., internal medicine, infectious disease) for easier monitoring of patients.[37]

NEED FOR A RANDOMIZED CONTROL TRIAL

Clinicians, patients, and society increasingly expect clinicians to practice evidence-based medicine, however their ability to do so is limited by the quality of existing evidence. In medicine, the randomized controlled trial (RCT) remains the gold standard because of the ability to control for measured and unmeasured confounders via random treatment allocation. To date the majority of RCTs assessing the effects of setting for stroke patients have been limited to Europe. [38,39] These studies identified that there was no significant difference in physiologic and activity functional outcomes for rehabilitation care that was provided in hospitals compared to patients homes.[38] Another stroke setting based RCT found that care in specialized stroke units improved outcomes compared to stroke care in general hospital wards.[39] However, because of differences in organizational structures between the systems the results from these trials may not be directly applicable to the United States.[4]

The value of a high-quality estimate of the comparative effectiveness of the two rehabilitation settings (i.e. IRF and SNF) generated from an RCT is primarily two-fold. For patients, it is important to identify which setting may improve a patient’s chance of achieving “maximum benefit”. Maximum benefit can be defined as the peak level of function and the time taken to achieve this peak function. For patients, steeper recovery trajectories may hasten the pace at which they are safely able to be discharged home, which is a patient centered outcome. Entwined within this is the fact that for stroke patients there is a limited time window (i.e. ~3 months) in which patients achieve >90% of their maximal physiologic and activity level functional gains.[11,40] This makes the initial discharge decision critical as there is limited opportunity to switch treatment.[36] Thus, if IRF care (compared to SNF care) improves outcomes then an underuse of IRF care will deprive stroke patients the opportunity the receive optimal care to ensure maximum odds of success.[41]

For health care purchasers (i.e. patients, the government, private insurers) there is widespread recognition of the need to pay for value not volume of care. Cost-effectiveness estimates that compare the outcomes and costs of care (i.e. value=outcomes/cost) are a powerful metric to empirically weigh tradeoffs between alternate approaches.[42,43] For CMS, the design and implementation of payment policies remains their most effective tool to drive changes in clinical practice.[30] However, the design of such policies are fraught with numerous challenges (i.e. self-interest, professional biases) and are limited by the amount and quality of data on treatment efficacy and costs that are needed to calculate valid cost effectiveness estimates.[8,30] Thus, changes to payment policies should be informed by high-quality, unbiased direct comparative effectiveness estimates which can identify the optimal patient level outcomes at the lowest cost.[30] For this dissertation, we will focus on the comparative effectiveness estimate and not costs as previous studies have identified that the total direct costs for IRF care are approximately double SNF care.[10,44]

For a trial that compares stroke rehabilitation at IRFs vs. SNFs, the nature of the two settings ultimately leads to the question of the effects of more intensive therapy over a short time with substantial clinical monitoring (IRF), compared to less intense therapy over the longer time period (SNF) with less stringent clinical monitoring.[6,7,45] Despite clear differences in indications, clinical oversight, and intensity of care between IRFs and SNFs, previous studies have stated that the real world evidence suggests that such a trial may be ethically justified for two reasons.[8,18]

First, there is a large overlap between IRF and SNF patients which provide exchangeable populations.[15] Second, IRF and SNF care is often thought to be interchangeable by many clinicains.[7,8,46] The interchangeable aspect is due to clinical uncertainty towards which patients should receive which type of care and an unclear comparative effectiveness estimate for IRF versus SNF care.[46]

The design of such an RCT is complicated by a multitude of logistical (i.e. identifying hospitals/facilities), practical (i.e. cost, patient enrollment), ethical (i.e. convincing hospitals and physicians to randomize patients), and measurement (i.e. within setting heterogeneity of therapy type, intensity etc.) issues.[8,47] A key complication for comparative effectiveness estimates for PAC, is that in the United States PAC is highly fragmented and there are large variations in therapy modalities, frequency of activities, and the quality of care that is delivered both between and within IRF and SNFs.[7,8] Thus, the specific relative effect for usual rehabilitation therapy (i.e., IRF or SNF) cannot be well quantified. Given the broad array of complexities involved in trial design, it is imperative for any trialist to conduct feasibility studies and model outcomes to carefully consider how to navigate these factors to improve the odds of trial success.[48]

One method to navigate these complexities, improve trial design and model the anticipated outcomes is to use a large observational database to design and ultimately emulate the desired RCT. Trial emulation is a form of observational data analysis which is guided by the principles of trial design (i.e. stated eligibility criteria, treatment strategies, treatment assignment, follow-up period, outcome, causal contrast, and statistical analysis).[49] In emulated trials, random treatment allocation is often emulated using propensity scores.[49,50] Propensity scores use existing covariates to estimate a patient’s probability of receiving treatment (versus control). Investigators can create two equivalent groups by using these estimates to either match or reweight patients (i.e. inverse probability of treatment weights).[51] Trial emulation is separate from traditional propensity score analysis because of the need to clearly define a time zero (t_0) at the time of randomization and a target trial protocol is developed based on the desired RCT.[49] Emulated trials are relatively new, but examples include using administrative claims data to inform the optimal timing of colon cancer screening,[52] data from large cohort studies to assess antiretroviral treatment switching strategies,[53] and the effect of postmenopausal hormone therapy on coronary heart disease.[54]

There are several advantages to employing an emulated trial framework to improve the design for a proposed RCT. First, investigators can explore the effect that various inclusion/exclusion criteria may have on eligible patients and facilities. Although RCT's remain the gold standard they often have poor external validity because they frequently take place at unrepresentative facilities (i.e. large academic hospitals). Recently there has been an increased focus to employ more pragmatic RCTs. Pragmatic RCTs randomize patients to real world clinical practice and include a broad range of eligible facilities.[55] These trials often have strong external validity, but weaker internal validity because of less stringent treatment protocols. Testing the effects of various patient and facility level inclusion/exclusion criteria can optimize this internal-external validity balance.[56,57] Second, pre-specifying the target trial protocol ensures that investigators are forced to test clinically meaningful interventions with easier clinical interpretation of the results.[49,58] Third, an explicit definition of time-zero (i.e. when eligibility criteria are met, treatment assignment, and outcomes counted) eliminates the risk of immortal time bias.[59] Immortal time bias is a form of selection bias, attributable to survivors having a longer time to receive more of a given exposure. This can lead to inaccurate conclusions of a protective effect for a given exposure.[59] Finally, emulated trials explore etiology under the counterfactual framework which resemble that of an RCT.[49,58]

OVERALL OBJECTIVES AND SPECIFIC AIMS

Given the substantial morbidity associated with stroke, and the value of alternative rehabilitation setting when applied to appropriately selected patients, there is a pressing need to identify which rehabilitation setting provides the best outcomes and value for stroke patients. Thus, the overarching purpose of this dissertation is to inform the design of a pragmatic trial to assess the relative effectiveness of rehabilitation at IRFs vs. SNFs for acute stroke patients, to assess its feasibility and to emulate it using observational data. Specifically, we will achieve this goal though a series of three specific aims:

Specific Aim 1: Use nationally representative administrative data to develop a multi-level multivariable logistic regression model to predict an acute stroke patient's probability of IRF versus SNF discharge.

Specific aim 1a: Identify patient-level, hospital-level and geographic predictors of IRF (versus SNF) discharge.

Specific aim 1b: Evaluate general and specific hospital contextual effects

Specific aim 1c: Characterize the heterogeneity of hospital effects on individual predicted probabilities of IRF (versus SNF) discharge.

Specific Aim 2: Identify a target trial population that will afford an optimal pragmatic-explanatory balance for a randomized control trial comparing the effectiveness of stroke rehabilitation care between IRFs and SNFs.

Specific aim 2a: Identify the effects that a stepwise application of various hospital level inclusion criteria has on characteristics and numbers of eligible hospitals, patients, and rehabilitation facilities.

Specific aim 2b: Assess trial generalizability by comparing target trial patients and hospitals to the national sample of all acute Medicare stroke patients who were discharged to an IRF or SNF.

Specific Aim 3: Use nationally representative administrative data to emulate three pragmatic clinical trials which compare the effectiveness of stroke rehabilitation at IRFs compared to SNFs.

Aim 3a. Determine the effect that greater rehabilitation facility level restrictions have on the comparative effectiveness of stroke rehabilitation at IRFs compared to SNFs.

CHAPTER 2: CHARACTERIZING HETEROGENEITY OF HOSPITAL EFFECTS FOR ACUTE STROKE PATIENTS PROBABILITY OF RECEIVING DISCHARGE TO AN INPATIENT REHABILITATION FACILITY OR A SKILLED NURSING FACILITIY

BACKGROUND

Every year, approximately 800,000 people in the United States experience a stroke and after an acute hospital stay of a few days, around half of these patients will be discharged to receive rehabilitation care at either an IRF or a SNF.[5,6] Both settings have the capacity to continue to medically assist patients, but their primary focus is on restoring physiological and activity level function to promote independent living.[4] In general, patients discharged to an IRF are expected to have significant physiological and activity level functional recovery gains leading to discharge back to the community. For around 2-3 weeks, these patients will receive time-intensive rehabilitation therapy (i.e., 3 hours a day) under direct physician oversight.[4] Conversely, patients discharged to a SNF generally are either unable to tolerate intensive therapy or have expectations of only moderate functional recovery. These patients will receive a broader range of more moderately intensive rehabilitation therapies under a physician monitored treatment plan for around 4-5 weeks.[4,10]

Despite different clinical indications for the two settings, nationally representative data shows that there is striking regional variation in discharge patterns to IRF and SNF facilities.[15–17,41] A large nationally representative cohort study showed that among the 918 hospitals with more than 15 stroke patients, the proportion of patients discharged to IRF versus SNF care ranged from 0-100%.[15] There are large cost implications for this variation, as the total direct medical costs for IRF care is approximately double that for SNF care for the six months preceding the stroke.[10,44] Several recent reports have outlined that addressing regional variation of post-acute care use is key to addressing variation CMS spending. IRFs and SNFs are the two most common inpatient settings of post-acute care and their use has increased over the past few decades because of shorter acute hospital LOS and an aging population.[24,5]

Several patient level sociodemographic factors (e.g., younger age, male sex, and having health insurance) and clinical factors (e.g., not having dementia, higher post-stroke physiological and activity level function, and fewer comorbidities) have all been found to be associated with IRF (vs. SNF) discharge.[15–17,41] However, beyond individual level factors, the context (e.g., the hospital, neighborhood, etc.) in which a patient received care at has a very large influence on the types of rehabilitation that patients receive.[60,61] For stroke patients, previous studies have identified that acute care hospitals have very large general contextual effects-as hospitals contribute around 30-50% of the overall variation in IRF and SNF discharge.[15–17] In addition, several studies have identified the role of several specific hospital contextual effects (i.e., specific associations between hospital level characteristics and discharge setting) and identified that hospital level factors such as for-profit status, having an affiliated IRF unit, and urban settings were all strongly associated with discharge setting.[15–17]

Previous studies all took a conventional quantitative epidemiological approach to explore specific patient- and hospital- level drivers of variation in IRF and SNF use.[15–17] This approach involves using a multivariable model to analyze individual fixed effect variables, with their interpretation being that variables were analyzed one at a time (i.e. the effect size is conditional on holding all other factors constant) and the magnitude of effect is reported as the differences between group average associations (i.e. adjusted odds ratios (aOR)).[62,63] Unfortunately, focusing on average effects ignores individual heterogeneity of patient level responses and the one-by-one interpretation of the strength of the associations between factors and discharge setting ignores the multitude of complex interactions that can occur between patient- and hospital-level factors.[63,64]

The transition towards more personalized healthcare has led to the development of several new methods that embrace individual heterogeneity. These methods account for individual heterogeneity of responses to either treatments or environments[62,65] We aimed to apply these methods to improve the understanding of hospital variation in IRF and SNF use by 1) Developing a prediction model to identify patient, hospital, and geographic predictors of IRF (vs. SNF) discharge; 2) Assessing general and specific hospital contextual effects; and 3) Evaluating heterogeneity of hospital effects on individual predicted probabilities of IRF (vs. SNF) discharge.

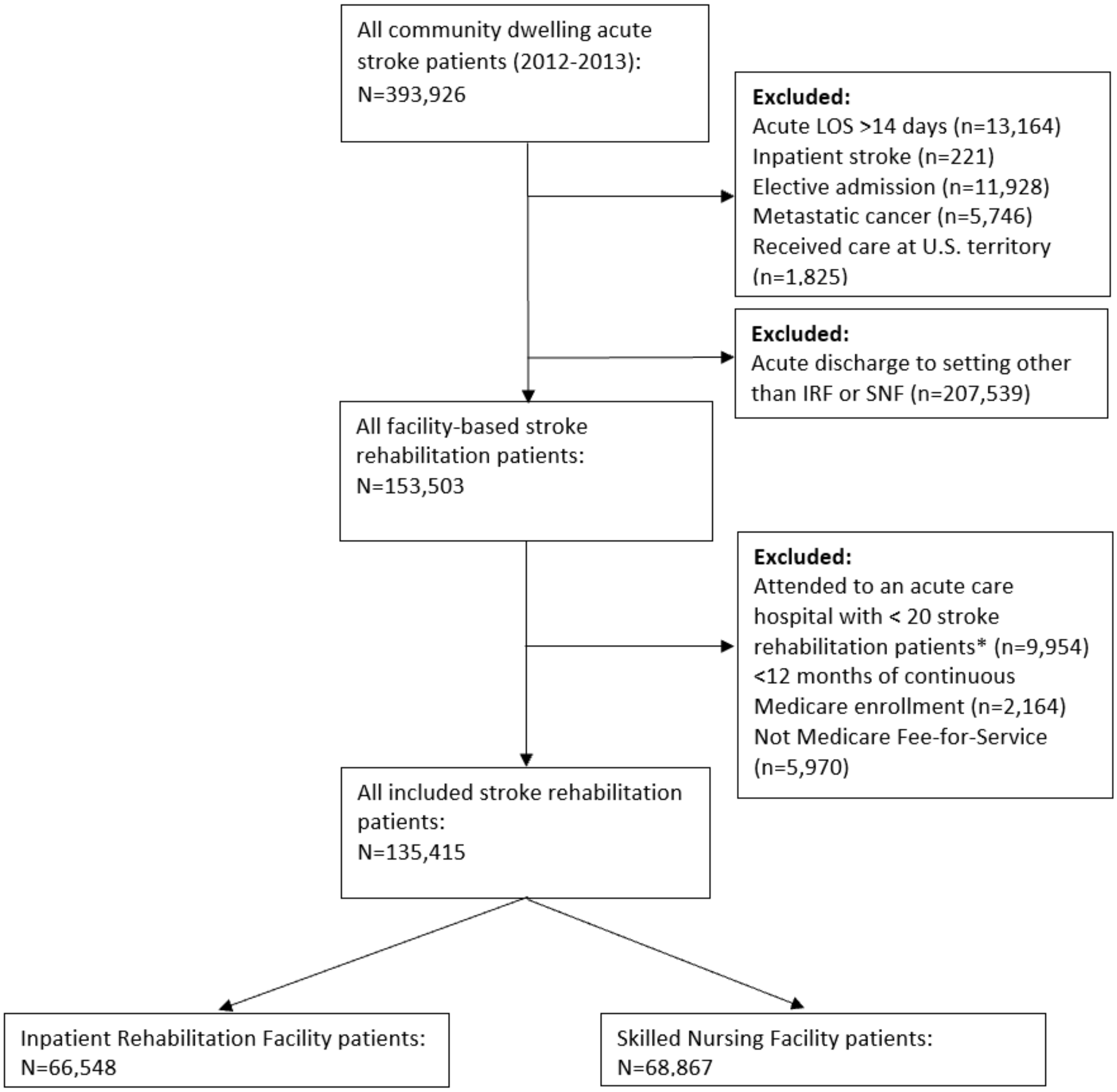
METHODS

Outcome

Our primary outcome was IRF vs. SNF discharge after hospitalization for acute stroke care. IRF and SNF patients were identified as patients who were discharged directly to an IRF or SNF and/or who subsequently were admitted to an IRF or SNF within 4 days of hospital discharge. Patients discharged to IRF and SNFs were identified based on hospital discharge code 62 and 03, respectively.

Stroke patients

We used Medicare standard analytic files from a 4-year period (2011-2014) to generate a retrospective cohort of community dwelling Medicare fee-for-service ischemic stroke or intracerebral hemorrhagic stroke patients with primary *International Classification of Diseases*, Ninth Revision (ICD-9), diagnosis codes of 431, 433.x1, 434.x1 who were admitted to an acute care hospital in the US between the two-year period: January 1st, 2012 and December 31st, 2013.[6] From the starting sample of 393,926 patients who were treated at 3,069 hospitals, we excluded patients for the following reasons: 1) Patients with an acute LOS > 14 days (n=13,164); 2) Inpatient stroke (n=221); 3) Elective admission (n=11,928); 4) Current diagnosis of metastatic cancer (n=5,746); 5) Received care in a U.S territory (n=1,825); 6) Discharged to a setting other than IRF or SNF (n=207,539); 7) Treated at an acute care hospital with <20 stroke patients discharged to either IRF or SNF rehabilitation setting in the 2-year window (n=9,954 and 1,223 hospitals); and 8) Not part of Medicare Fee-for-service (n=5,970). We excluded smaller hospitals to ensure more accurate random intercepts in our subsequent multi-level models. The resulting sample comprised 135,415 patients who were treated at 1,816 hospitals. Figure 2.1 shows the study flow diagram for how the final retrospective cohort was assembled.



Abbreviations: LOS: Length of stay
*Stroke rehabilitation patients: Discharged to an Inpatient Rehabilitation Facility or a Skilled Nursing Facility

Figure 2.1: Flow diagram describing the generation of the final study cohort for Aim 1

Data sources

The analytic dataset was comprised of the following Medicare administrative files (the details of which are shown in Supplemental Table 2.1): Inpatient Claims (IPC),[66] the inpatient and SNF Medicare Provider Analysis and Review (MedPAR) files,[66] Part B Carrier Summary

Data File (Part B file),[66] the Master Beneficiary Summary File (MBSF),[66] the American Community Survey (ACS),[67] the Provider of Service File (POS),[68] and Medicare's Hospital Compare database.[69] We included data from 2011 until 2014 to allow at least 1 year of information on pre-stroke function/health and at least 1 year of follow-up. The IPC file provided information on ICD-9 diagnosis (including the indexed stroke) and procedure codes, as well as identified if the patient was treated at a Hospital, IRF, or SNF. MedPAR provided aggregated information for a single stay and categorized in-hospital charges. The MBSF provided information on age, race, sex, enrollment reason, zip code, and disability information from social security. The Part B file was used to identify Current Procedural Terminology codes (CPT) for physical therapy (PT), occupational therapy (OT), and speech language-pathology (SLP) provided during the acute in-patient stay. The ACS provided race and sex specific zip code level aggregate data for information on income and educational attainment. The POS file provided information on hospital characteristics and Medicare's Hospital Compare data provided information on hospital quality by providing patient case-mix adjusted measures of hospital processes and outcomes. Combining these files enabled us to capture all claims at both the acute and rehabilitation facility level. Files were linked using Medicare beneficiary identifiers (for patients) or hospital provider number (for hospitals/rehabilitation facilities).

Covariates

A comprehensive list of all patient factors (hypothesized to potentially influence IRF or SNF discharge after stroke) used in this study along with their technical definitions (i.e. their ICD-9 codes) can be found in Supplemental Table 2.2. <u>Demographic covariates</u> included age, sex, and race (white, black, Hispanic, and other). Measures of <u>prior health care utilization</u> were taken 1 year prior to the indexed stroke event and included; the number of hospitalizations, home-time (i.e. time alive and at home (i.e., not in an acute care hospital, IRF or SNF),[70] previous use of IRF (yes/no), and previous use of SNF (yes/no). Clinical information included the Elixhauser Comorbidity Index (which consists of 31 comorbidities) and any dementia documented during the indexed hospitalization.[71] We obtained information during the time of the indexed stroke because we did not have access to other measures of comorbidities that collect data on comorbidities in the year prior (e.g., CMS's Hierarchical Condition Categories). Available stroke-related information was collected during the index hospitalization which included stroke subtype (ischemic or intracerebral hemorrhagic) and stroke severity (mild, moderate, severe). Stroke severity was categorized using the stroke administrative severity index.[72] This index is comprised of five ICD-9 discharge diagnostic stroke symptoms (i.e. aphasia, coma, dysarthria/dysphagia, hemiplegia/monoplegia, and neglect) and two ICD-9 procedure codes (i.e. parenteral infusion and tracheostomy/ventilation) which were weighted based on the strength of their association with 30 day mortality.[72] This index has been shown to be strongly correlated with the NIH Stroke Scale in Medicare patients.[72] In addition, we used several hospital health services measures as proxies for medical acuity which have previously been shown to be strongly correlated with medical acuity. These included LOS, the number of days spent in the intensive care unit (ICU days) and the number of days spent in the coronary care unit (CCU days), emergency department (ED) admission (based on any ED charge data), and six lifesaving procedures (i.e., hemodialysis, gastrostomy tube, intubation/ventilation, cardiopulmonary resuscitation, enteral or parenteral, and tissue plasminogen activator (tPA) use).

Hospital level variables included the number of hospital beds (per 50 bed increase), medical school affiliation (yes/no), hospital ownership (church, private-not for profit, private-for profit, government, other), whether the hospital had an IRF unit directly associated with it (yes/no), whether the hospital was classified as urban or rural, and finally, the 10 CMS regions (further details available in Supplemental Table 2.2)

Analysis

Distribution of patient- and hospital-level factors for both IRF and SNF populations was described using means and standard deviations for continuous variables and percentages for categorical variables. For binary level comparisons- specifically comparing characteristics between patients that were discharged to an IRF or SNF we used absolute standardized differences (ASDs) rather than traditional statistically significance testing (p-values) because ASDs are not affected by the large sample size. We considered ASDs greater than 0.1 to be clinically meaningful.[73] For continuous variables ASDs were calculated using the formula $ASD = \frac{|X_{IRF} - X_{SNF}|}{\sqrt{\frac{S_{IRF}^2 + S_{SNF}^2}{2}}}$. Where $X_{IRF} - X_{SNF}$ is the difference in the sample mean of IRF and SNF patients, and S_{IRF}^2 and S_{SNF}^2 are the sample variances for IRF and SNF patients. For categorical variables ASDs were calculated using the formula $ASD = \frac{|p_{IRF} - p_{SNF}|}{\sqrt{\frac{p_{IRF}(1-p_{IRF}) + p_{SNF}(1-p_{SNF})}{2}}}$.[73] where $p_{IRF} - p_{SNF}$ is the difference in the prevalence of the covariate in the IRF and SNF populations respectively.[73] For three-way comparisons, chi-square and one-way ANOVA was used to test for statistical significance which was set at p<0.01.

Prediction model development

For prediction model development, we followed the recommendations from the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.[74] Single level univariate analyses were first used to compute crude odds ratios (ORs) for the association between all covariates and discharge to IRF (vs.

SNF). We then developed a multivariable single level logistic regression model (Model 1) to predict the probability of IRF discharge (vs. SNF). Because of the large sample size, we were not concerned with overparameterizing the model and all 80 covariates listed in Supplemental Table 2.2 were included. This single level prediction model was developed as follows. First, the data were randomly split into derivation (n=81,249, 60%) and validation (n=54,166, 40%) samples. Second, within the derivation sample we optimized the functional form (i.e. improved any non-linear fit between predictors and the outcome) for seven continuous predictor variables (i.e. age, Elixhauser comorbidity index, pre-stroke home-time, number of pre-stroke hospitalizations, hospital LOS, ICU days , and CCU days) using fractional polynomials.[75] Fractional polynomials optimizes the functional form by testing a series of power transformations (e.g., x^{-2} or x^3) in up to two terms (e.g., x^{-2} *and* x^3) for each continuous variable. Models with alternate parameterizations were compared using the likelihood ratio test. Models with fewer terms were nested in larger models and the smallest statistically insignificant model was selected. Third, we tested a set of two-way interactions using a significance level of ($p<0.01$). The interaction set comprised all two-way combinations of the following eight a-priori selected variables (i.e. age, sex, race, LOS, Elixhauser comorbidity index, dementia, stroke subtype, and pre-stroke SNF use) that were expected to influence discharge setting. These a-priori variables were selected based on prior literature and clinical plausibility.[15,16,41,76] Interaction significance was set at $p<0.01$ – this conservative threshold was chosen because there were 28 interactions tested and we had a very large sample size.

Following the development of the single level prediction model (Model 1), we then introduced a hospital-specific random effect (RE) term into the model (Model 2). This model was a hierarchical logistic regression model with patient level fixed effects and a hospital R.E.

The data structure was multilevel because patients were nested within hospitals. This nesting structure accounts for clustered observations within hospitals and allows for the partitioning of variances of individual patients and the total variance. Finally, in the fully adjusted model (model 3) all available hospital level variables (i.e., bed size, hospital ownership, urban setting) were added as fixed effects to model 2.

We evaluated the performance of each model by assessing model discrimination and model calibration. Model discrimination was evaluated using the C statistic. The C statistic indicates the probability that among two randomly selected patients – one who had an event (i.e., IRF discharge) and one who did not have an event (i.e., SNF discharge), the one who experienced the event had a higher predicted risk.[75] A C statistic of 0.5 indicates a model is no better than random chance and 1.0 indicates perfect prediction.[77] Model calibration was evaluated using Hosmer-Lemeshow goodness-of-fit tests and calibration plots. Calibration plots, show the predicted risk for IRF (vs. SNF) discharge over 10 deciles of predicted risk.[75,77] Well calibrated models have slopes close to 1.[77]

Estimating general and specific hospital contextual effects

We assessed general and specific hospital contextual effects by following a previously developed multistep framework.[62] General contextual effects are a reflection of the degree to which the context (i.e. the specific cluster of interest-which in this case is the hospital) influences patient level outcomes and were measured by the change in the C-statistic and the Intraclass Correlation Coefficient (ICC) (see steps 1 and 2 in the following paragraph).[60,62,78] Conversely, specific contextual effects (see steps 3 and 4) measure the associations of specific contextual factors (i.e. hospital level characteristics) and patient level outcomes and were measured by the proportional change in variance (PVC) and the 80% Interval Odds Ratio (IOR).[62,78] General and specific hospital contextual effects are related, in that when general hospital contextual effects are large, then specific hospital contextual effects have less precision.[78]

General contextual effects

In Step 1 we calculated the change in C statistics between the single level patient prediction model and multi-level patient prediction model-which also included the hospital RE (i.e., C-statistic $_{\text{model 2}}$ - C-statistic $_{\text{model 1}}$). The change in C-statistic is an estimate of how much additional predictive value the hospital RE added, with a larger change indicating a larger general hospital contextual effect.[62,78] We also calculated the change in the C statistics between C-statistic $_{\text{model 3}}$ - C-statistic $_{\text{model 1}}$ to identify the additional predictive value of accounting for specific hospital characteristics (i.e., specific contextual effects). In Step 2 we calculated unadjusted and adjusted ICCs for models 2 and 3. The ICC measures the proportion of total individual level variation in IRF and SNF discharge which can be attributed to variation between hospital random intercepts.[78,79] The larger this proportion is, the larger the general contextual hospital effect is estimated to be.[78] The ICC was calculated using the equation: $ICC = \sigma^2 / (\sigma^2 + \frac{\pi^2}{3})$. Where σ^2 is the variation of the hospital random intercepts.

Specific contextual effects

In Step 3 we calculated the PVC, which measures the proportion of hospital level variation in IRF (vs. SNF) discharge which could be explained by the addition of hospital characteristics (i.e. the specific hospital contextual factors) which were added to model 3. The PVC was calculated using the equation: $PVC = \frac{\sigma^2\text{ (model 2)} - \sigma^2\text{ (model 3)}}{\sigma^2\text{(model 2)}}$ where σ^2 is the variation of hospital random intercepts estimated from the multi-level model.[79] Finally, in Step 4 we calculated the 80% IOR for all hospital level characteristics. The IOR is a reflection of the amount of hospital variation that is present for each hospital level characteristic (i.e. each specific contextual factor).[62,79] A wide IOR reflects that there is substantial variation in the effect of that specific hospital contextual factor (i.e., hospital characteristics). The width of the IOR will be larger when general contextual effects are large (i.e., when the variance of the hospital level RE is large).[80] The lower and upper 80% IORs were calculated using the equations: ($IOR_{lower} = \exp(\alpha + \sqrt{2}\sigma^2\Phi^{-1}(0.10)$, $IOR_{upper} = \exp(\alpha + \sqrt{2}\sigma\Phi^{-1}(0.90)$.[62,79] Where α is the regression coefficient with the hospital level covariate, σ^2 is the variation of hospital random intercepts and Φ are the percentiles from the normal standard deviation.

Heterogeneity of hospital effects on individual predicted probabilities for IRF (vs. SNF) discharge

We quantified the individual heterogeneity of responses to hospital effects by broadly following an approach previously developed to assess heterogeneity of treatment effects between individuals participating in RCTs.[81] This method assumes that the magnitude of the exposure effect (i.e. treatment) is confounded by the baseline risk.[81] Because we were interested in hospital effects, hospitals were stratified into SNF favoring, typical, and IRF favoring hospitals. We used a previously developed method to stratify hospitals by their propensity to discharge patients to IRFs or SNFs after adjustment for patient case mix.[82,83] Using model 2, each hospital was ranked based on the empirical mean Bayes estimate of the hospitals random intercept. This was estimated as the logarithm of the odds ratio of IRF (vs. SNF) discharge at each hospital and was compared to the average hospital (which has a RE of 0).[83] The 99% confidence intervals (CIs) of these random intercepts were then estimated using the standard error of the hospital RE term (σ^2). We used 99% CIs because previous studies have shown there is substantial hospital level variation in IRF (vs. SNF) discharge.[15–17] Hospitals with statistically significant ($p<0.01$) negative random intercepts were considered SNF favoring hospitals, and hospitals with statistically significant positive random intercepts were considered IRF favoring hospitals.[83] All other hospitals i.e., those with random intercepts that were not statistically significantly different from 0, were considered typical hospitals. The use of confidence intervals to classify hospitals has been shown to be a reliable method.[84,85]

Within each hospital type (i.e. SNF favoring, typical, IRF favoring), we then calculated the proportion of patients which had either substantial (i.e. >20%), considerable (i.e. >10%), or minimal (i.e. <10%) change in their predicted probabilities ($\hat{p}$) for IRF discharge (relative to model 1) following the additional adjustments made in models 2 (i.e., $\hat{p}_{\text{model 1}} - \hat{p}_{\text{model 2}}$) and 3 (i.e., $\hat{p}_{\text{model 1}} - \hat{p}_{\text{model 3}}$). The magnitude of patient level heterogeneity introduced by hospital effects is thus identified by the proportion of patients which had either considerable or substantial changes in their predicted probabilities following hospital adjustment.[86] Any change in a predicted probability of IRF discharge is directly impacted by the hospital that they went to, with small changes indicating a small hospital effect and large changes indicating a large hospital effect. Finally, we assessed general and specific hospital contextual effects by calculating ICCs, PVCs, and 80% IORs for each type of hospital.

RESULTS

The final sample included 135,415 patients, which were evenly split between IRF (n=66,548 49.1%) and SNF (n=68,867, 50.9%) patients. Details on how the final sample was attained is shown in the study flow diagram (Figure 2.1). Table 2.1 presents ASDs between the IRF and SNF populations for important baseline patient level factors. These factors were selected based on statistically significant p-values and clinical relevance. Based on ASD values of >0.10, meaningful differences between IRF and SNF populations indicate that IRF patients were younger, more likely to be male, had better pre-stroke function (i.e. more home-time, less hospitalizations in the year prior to the stroke event), and were more likely to receive tPA (Table 2.1). In contrast, SNF patients were more likely to have dementia, and a have received a gastrostomy tube.

Table 2.1: Differences in selected baseline patient characteristics for acute Medicare stroke survivors who were discharged to an IRF or SNF

	Whole sample (%) n=135,415	IRF patients (%) n=66,548	SNF patients (%) (n=68,867	ASD*
Demographic characteristics:				
Age in years (SD)	81.5 (8.0)	79.4 (7.7)	83.4 (7.9)	0.51
Race				
White	80.8	80.0	81.5	0.04
Black	11.4	11.6	11.2	0.01
Hispanic	4.5	4.9	4.2	0.03
Other	3.3	3.5	3.1	0.02
Female	61.1	56.2	68.5	0.20
Pre-stroke functional proxies:				
Days of previous home-time¶ (SD)	358.5 (21.1)	361.9 (11.6)	355.2 (26.8)	0.32
Previous hospitalization	20.5	15.8	25.0	0.23
Previous SNF use	11.3	4.8	17.7	0.42
Previous IRF use	2.7	3.3	2.2	0.07
Comorbidities:				
Total number of Elixhauser comorbidities	4.0 (1.8)	4.0 (1.8)	4.0 (1.9)	0.02
Dementia	9.1	4.4	13.6	0.33
Stroke characteristics:				
Stroke subtype				
Ischemic	91.0	90.9	91.0	0.01
Intracerebral hemorrhagic	9.0	9.1	9.0	0.01
Stroke severity				
Mild	39.1	39.6	38.5	0.02
Moderate	39.2	39.2	39.2	<0.01
Severe	21.7	22.2	21.2	0.02
Hospital Health Services Use				
Length of stay	5.1 (2.7)	5.1 (2.7)	5.2 (2.7)	0.02
ICU days	1.8 (2.6)	1.8 (2.6)	1.8 (2.7)	0.01
CCU days	0.6 (1.7)	0.7 (1.8)	0.6 (1.6)	0.05
ED admission	90.4	89.4	91.4	0.07
Lifesaving procedures				
Hemodialysis	1.3	1.1	1.6	0.04
Gastrostomy tube	6.1	3.9	8.2	0.18
CPR	0.1	0.1	0.1	<0.01
Parenteral nutrition	3.0	2.4	3.7	0.07
Intubation/ventilation	1.8	1.9	1.7	0.01
tPA	6.5	8.0	5.0	0.13

Abbreviations: ASD: Absolute Standardized difference, IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility, ED: Emergency department, ICU: Intensive care unit, CCU: Cardiology care unit, tPA: Tissue plasminogen activator *Absolute standardized differences >0.1 were considered significant.
¶ Home-time: Days spent alive and outside of the acute hospital, IRF or SNF

Table 2.2 shows selected baseline hospital level characteristics for all 1,816 hospitals which were included in the final sample. On average these hospitals had 341 beds, just under half (43.9%) were private-not for profit, half of the hospitals had an affiliation with an IRF unit, and most (84.1%) were situated in an urban setting. CMS regions 4 (AL, FL, GA, KY, MS, NC SC, TN) and 5 (IL, IN, MI, MN, OH, WI) were the two regions with the most hospitals.

Table 2.2: Selected baseline hospital characteristics among hospitals that treated Medicare stroke survivors who were discharged to an IRF or SNF (n=1,816 hospitals)

Hospital baseline characteristics	Mean or %
Bed count (SD)	341.3 (247.0)
Hospital ownership	
Church	13.2
Private-not for profit	43.9
Private-for profit	16.5
Government	6.6
Other	19.9
Medical school affiliation	42.8
Affiliated IRF unit	50.0
Urban setting (vs. rural)	84.1
CMS region	
1) CT, ME, MA, NH, RI, VT	5.8
2) NY, NJ	8.7
3) DE, DC, MD, PA, VA, WV	10.3
4) AL, FL, GA, KY, MS, NC SC, TN	21.8
5) IL, IN, MI, MN, OH, WI	18.8
6) AR, LA, NM, OK, TX	11.3
7) IA, KS, MO, NE	5.2
8) CO, MT, ND, SD, UT, WY	2.9
9) AZ, CA, HI, NV	11.6
10) AK, ID, OR, WA	3.6

Abbreviations: SD: Standard deviation, IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility, CMS: Center for Medicare & Medicaid Services
CMS regions include state abbreviations

Table 2.3 shows patient level differences in the ASDs between selected hospital characteristics. Thus, the average IRF patient was more likely to be treated at a large hospital, a hospital with an affiliated IRF unit, and at a hospital that was situated in an urban setting compared to the average SNF patient. There were two differences in CMS regions, with patients in CMS region 4 (AL, FL, GA, KY, MS, NC SC, TN) being more likely to go to a SNF, whereas patients in CMS region 6 (AR, LA, NM, OK, TX) were more likely to go to an IRF.

Table 2.3: Selected patient level differences in hospital characteristics among hospitals that treated Medicare stroke survivors who were discharged to an IRF or SNF (n=1,816 hospitals)

	Whole sample (n=135,415 [%])	IRF patients (n=66,548 [%])	SNF patients (n=68,867 [%])	Absolute Standardized Differences*
Bed count (SD)	461.0	479.1 (321.9)	443.6 (327.4)	0.11
Hospital ownership				
Church	14.6	14.7	14.4	0.01
Private-not for profit	47.5	47.1	48.0	0.02
Private-for profit	12.4	13.2	11.7	0.05
Government	5.5	5.4	5.5	0.01
Other	20.0	19.6	20.4	0.02
Medical school affiliation	52.7	54.9	50.6	0.09
Affiliated IRF unit	56.6	63.6	49.8	0.28
Urban setting (vs. rural)	91.0	92.8	89.2	0.13
CMS region				
1) CT, ME, MA, NH, RI, VT	5.9	6.4	5.4	0.04
2) NY, NJ	9.9	10.3	9.4	0.03
3) DE, DC, MD, PA, VA, WV	11.4	11.4	11.3	<0.01
4) AL, FL, GA, KY, MS, NC SC, TN	22.3	24.4	20.2	0.10
5) IL, IN, MI, MN, OH, WI	18.6	19.2	17.9	0.03
6) AR, LA, NM, OK, TX	11.2	8.2	14.4	0.20
7) IA, KS, MO, NE	5.2	4.3	6.0	0.08
8) CO, MT, ND, SD, UT, WY	2.2	1.9	2.5	0.04
9) AZ, CA, HI, NV	10.0	10.2	9.9	0.01
10) AK, ID, OR, WA	3.3	3.7	2.9	0.05

Abbreviations: CMS: Center for Medicare & Medicaid Services (state abbreviations), Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility. *Absolute standardized differences >0.1 were considered significant.

Patient- and hospital- level factor associations with IRF (vs. SNF) discharges

A full list of all patient and hospital level factors associations with discharge to an IRF vs. SNF from both the univariate analysis and from models 1, 2 and 3 is shown in Supplemental Table 2.3. For brevity, Table 2.4 presents associations of selected (based on clinical importance) patient level factors and IRF (vs. SNF) discharge. For ease of interpretation these associations do not include the interactions and fractional polynomials which were included in the final prediction models. The following summary of major findings are based on the fully adjusted model (i.e. model 3). Age (aOR): 0.93, 95%CI: 0.93, 0.93), female sex (aOR: 0.74, 95%CI: 0.71, 0.76), and black (vs. white) race (aOR: 0.90, 95%CI: 0.85, 0.94) were all statistically significantly associated with lower odds of discharge to an IRF (vs. SNF). For age, an aOR of 0.93 indicates that a 1-year increase in age was independently associated with 7% lower odds of IRF discharge. Several pre-stroke variables, that likely act as proxies for functional status had significant associations with IRF discharge. For example, the use of an IRF (aOR: 1.89, 1.73, 2.05) or SNF (aOR: 0.39, 95%CI: 0.37, 0.42) in the year prior to the index stroke both had particularly strong associations. Apart from dementia, (aOR: 0.30, 95%CI: 0.28, 0.31) clinical comorbidities generally did not have statistically significant associations with discharge setting. Patients who received hemodialysis (aOR: 0.67, 95%CI: 0.59, 0.75) or a gastrostomy tube (aOR: 0.42, 95%CI: 0.39, 0.45) had decreased odds of IRF (vs. SNF) discharge. Significant hospital level characteristics included for-profit status (aOR: 1.34, 95%CI 1.14, 1.58) and an affiliation with a medical school (aOR: 1.12, 95%CI: 1.02, 1.35). Not surprisingly, if a hospital had an affiliated IRF unit it increased the odds of IRF discharge by more than 2-fold (aOR: 2.53, 95%CI: 2.25, 2.84) as did urban setting (aOR: 1.71, 95%CI: 1.44, 2.03). There were also important geographic contextual effects between CMS regions. For example, patients in CMS region 6 (AR, LA, NM, OK, TX) vs. CMS region 5 (IL, IN, MI, MN, OH, WI) had over three times the odds of IRF discharge (aOR: 2.12, 95%CI: 2.54, 3.82).

Table 2.4: Unadjusted and adjusted odds ratio associations of selected patient and hospital contextual factors with IRF (vs. SNF) discharge among Medicare stroke survivors-multivariable logistic regression results

	Unadjusted (OR)	95 % CI	Model 1 (aOR)	95 % CI	Model 2 (aOR)	95 % CI	Model 3 (aOR)	95 % CI
Demographic characteristics								
Age	0.94	[0.97-0.99]	0.94	[0.94,0.94]	0.93	[0.93,0.93]	0.93	[0.93,0.93]
Race (ref=White)								
Black	1.05	[1.02,1,09]	0.94	[0.90,0.98]	0.90	[0.86,0.95]	0.90	[0.85,0.94]
Hispanic	1.19	[1.13,1.26]	1.13	[1.07,1.20]	0.98	[0.91,1.05]	0.96	[0.89,1.03]
Other	1.16	[1.09,1.23]	0.98	[0.92,1.04]	1.04	[0.96,1.12]	1.03	[0.95,1.11]
Female sex	1.23	[1.23,1.26]	0.77	[0.75,0.79]	0.74	[0.71,0.76]	0.74	[0.71,0.76]
Pre-stroke functional proxies								
Pre-home-time	1.02	[1.02, 1.03]	1.01	[1.01,1.01]	1.01	[1.01,1.01]	1.01	[1.01,1.01]
Previous number of hospitalizations	0.72	[0.71,0.73]	0.92	[0.90,0.94]	0.92	[0.90,0.94]	0.92	[0.90,0.94]
Previous SNF use	0.23	[0.23,0.24]	0.39	[0.37,0.42]	0.39	[0.36,0.42]	0.39	[0.37,0.42]
Previous IRF use	1.58	[1.48,1.69]	2.85	[2.63,3.08]	1.89	[1.73,2.05]	1.88	[1.73,2.05]
Comorbidities								
Total Elixhauser comorbidity index (0-31)	0.99	[0.98,0.1.0]	0.98	[0.94,1.01]	0.98	[0.94,1.02]	0.98	[0.94,1.02]
Dementia	0.290	[0.28,0.33]	0.35	[0.33,0.37]	0.3.0	[0.28,0.31]	0.3.0	[0.28,0.31]
Stroke characteristics:								
Stroke subtype (ref=Ischemic)								
Intracerebral hemorrhagic	1.02	[0.98,1.06]	1.00	[0.96,1.04]	1.00	[0.96,1.05]	1.00	[0.96,1.05]
Stroke severity (ref=mild)								
Moderate	1.03	[1.00,1.05]	1.01	[0.96,1.06]	1.00	[0.95,1.06]	1.01	[0.96,1.06]
Severe	1.08	[1.05,1.11]	1.01	[0.93,1.10]	1.00	[0.91,1.01]	1.0010	[0.91,1.10]

Table 2.4 (cont'd)

Hospital health service use								
Length of stay (days)	0.99	[0.99,1.00]	0.99	[0.98,0.99]	0.99	[0.99,1.00]	0.99	[0.99,1.00]
ICU days (per 1-day increase)	1.00	[1.00-1.00]	1.00	[0.99,1.00]	0.99	[0.98,1.00]	0.99	[0.98,1.00]
CCU days (per 1-day increase)	0.97	[0.96, 0.98]	0.98	[0.97, 0.98]	0.96	[0.95, 0.97]	0.96	[0.95, 0.97]
EMS admit	0.80	[0.77,0.83]	0.86	[0.82,0.89]	0.97	[0.92,1.02]	0.97	[0.93,1.02]
Lifesaving procedures								
Hemodialysis	0.70	[0.64,0.77]	0.71	[0.64,0.79]	0.67	[0.59,0.75]	0.67	[0.59,0.75]
Gastrostomy tube	0.45	[0.43,0.47]	0.47	[0.44,0.50]	0.42	[0.39,0.45]	0.42	[0.39,0.45]
CPR	0.97	[0.60,1.59]	1.13	[0.66,1.95]	1.34	[0.74,2.42]	1.35	[0.75,2.43]
Parenteral Nutrition	0.65	[0.61,0.69]	0.97	[0.90,1.05]	1.08	[1.00,1.18]	1.09	[1.00,1.19]
Intubation /ventilation	1.10	[1.01,1.19]	1.04	[0.95,1.15]	1.20	[1.08,1.33]	1.20	[1.09,1.33]
tPA	1.68	[1.60,1.75]	1.82	[1.72,1.92]	2.09	[1.97,2.22]	2.10	[1.97,2.23]
Hospital Characteristics								
Bed count per 50 bed increase	1.02	[1.02,1.02]					1.00	[0.99,1.01]
Hospital ownership (ref=private not for profit)								
Church	1.04	[1.00,1.7]					1.09	[0.92,1.29]
Private-for profit	1.15	[1.11,1.19]					1.34	[1.14,1.58]
Government	1.00	[0.95,1.05]					0.81	[0.64,1.02]
Other	0.98	[0.95,1.01]					0.97	[0.84,1.12]
Medical school affiliation	1.19	[1.16,1.22]					1.17	[1.02,1.35]
IRF unit	1.76	[1.72,1.80]					2.53	[2.25,2.84]
Urban (vs. rural)	1.56	[1.51,1.62]					1.71	[1.44,2.03]

Table 2.4 (cont'd)

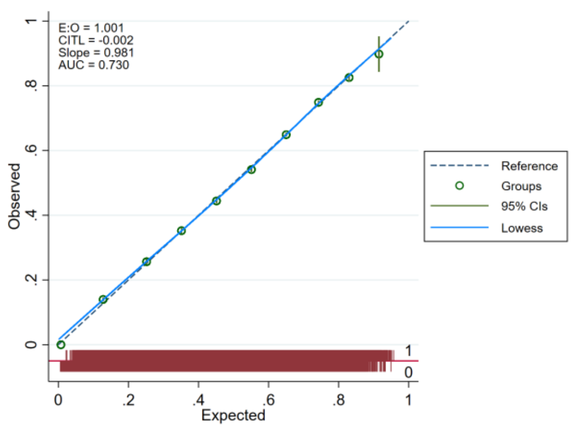
CMS region (ref= 5 (IL, IN, MI, MN, OH, WI)						
1) CT, ME, MA, NH, RI, VT	0.91	[0.87,0.96]			1.05	[0.81,1.36]
2) NY, NJ	0.98	[0.94,1.02]			1.39	[1.11,1.73]
3) DE, DC, MD, PA, VA, WV	1.07	[1.02,1.11]			1.24	[1.01,1.53]
4) AL, FL, GA, KY, MS, NC SC, TN	0.89	[0.86,0.92]			1.30	[1.09,1.55]
6) AR, LA, NM, OK, TX	1.88	[1.81,1.96]			3.12	[2.55,3.83]
7) IA, KS, MO, NE	1.49	[1.41,1.57]			1.69	[1.30,2.20]
8) CO, MT, ND, SD, UT, WY	1.37	[1.27,1.47]			0.99	[0.71,1.38]
9) AZ, CA, HI, NV	1.04	[1.00,1.09]			1.35	[1.10,1.66]
10) AK, ID, OR, WA	0.82	[0.77,0.88]			0.68	[0.50,0.92]

Abbreviations: IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility, ED: Emergency department, ICU: Intensive care unit, CCU: Cardiology care unit, tPA: Tissue plasminogen activator, CMS: Center for Medicare & Medicaid Services, OR: Odds Ratio, aOR: Adjusted Odds Ratio
Model 1: Single level logistic regression model that modeled discharge to an IRF vs. SNF that included patient level fixed effects
Model 2: Multi-level logistic regression model that modeled discharge to an IRF vs. SNF that included patient level fixed effects and a hospital random effect
Model 3: Multi-level logistic regression model that modeled discharge to an IRF vs. SNF that included patient and hospital level fixed effects and a hospital random effect
CMS regions include state abbreviations

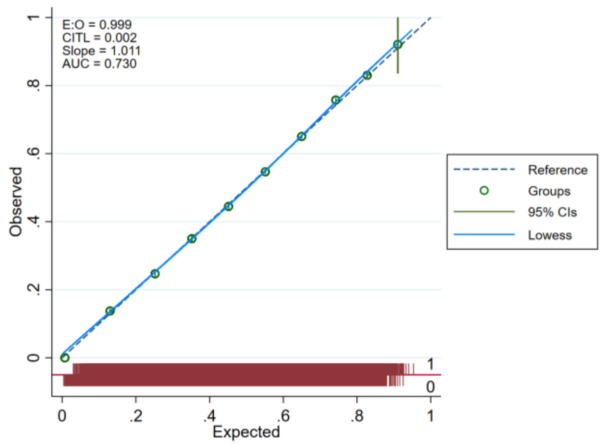
Patient level prediction model for IRF (vs. SNF) discharge

The developed single patient level prediction model (model 1) had non-linear functional forms for age, LOS, ICU days, and CCU days and included three statistically significant ($p<0.01$) 2-way interactions (i.e. age*pre-stroke SNF use, race*sex, and dementia*pre-stroke SNF use). The derivation (C-Statistic=0.73) and validation (C-statistic=0.73) results had similar discrimination and calibration (Figure 2.2). Models with C statistics above 0.7 are considered to have "good" discrimination while both calibration plots had slopes close to 1.[77] The Hosmer-Lemeshow goodness of fit tests were statistically significant for both the derivation (p=0.002) or validation (p=0.044) samples which may indicate a poor model fit. However, this test has been shown to be extremely sensitive to large samples sizes like we have in this dataset.[87] Because the model fit for the deviation and validation data sets were very similar, and the betas were almost identical we proceeded by fitting models from the entire dataset.

Derviation Sample:



Validation Sample:

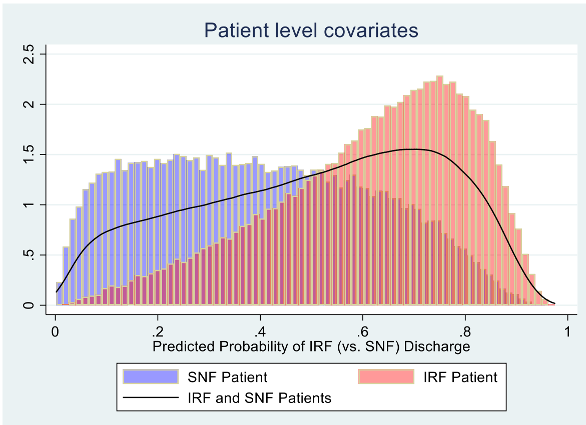


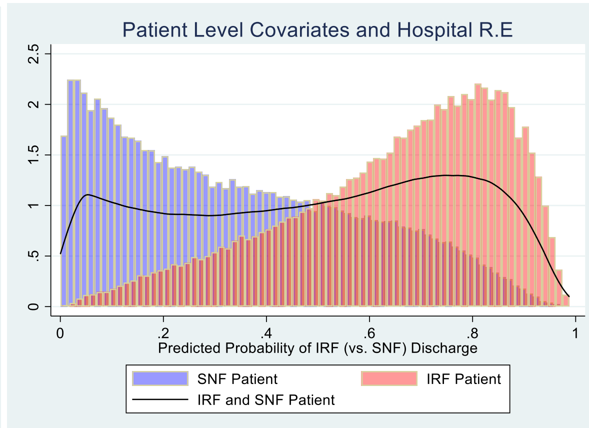
Abbreviations:
E:O: Expected to observed, CITL: Calibration in the large, AUC: Area under the curve, IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility
*Note: The AUC is the same measure as the C-statistic

Figure 2.2: Calibration plots for the derivation and validation samples from the single level multivariable logistic regression model (model 1) that predicted IRF or SNF discharge for acute Medicare stroke patients

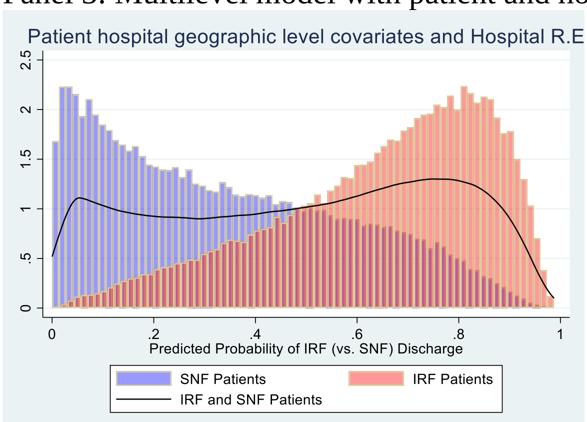
The distribution of predicted probabilities for IRF and SNF patients, along with a kernel density estimate of the predicted probability for the whole sample from model 1 is shown in the first panel of Figure 2.3. Adding the hospital RE (model 2) substantially increased model discrimination (C-Statistic=0.82) and resulted in a dramatic shift in the predicted probability distributions (Figure 2.3, panel 2). The shift was primarily due to many more patients having very low (i.e. <20%) predicted probabilities once the hospital RE was accounted for in the model. Further adjustment by adding hospital level characteristics (model 3) did not result in a meaningful change for either model discrimination (C-Statistic=0.82) or the predicted probability distributions (Figure 2.3, panel 3).

Panel 1: Single level model with patient factors   Panel 2: Multilevel model with patient factors





Panel 3: Multilevel model with patient and hospital factors



Panel 1: Single level logistic regression model with patient level fixed effects [C-statistic=0.73 (model 1)]
Panel 2: Hierarchical logistic regression model with patient level fixed effects and hospital random effect [C-statistic=0.82 (model 2)]
Panel 3: Hierarchical logistic regression model with patient, and hospital contextual factors as fixed effects and hospital random effect [C-statistic=0.82 (model 3)]

Figure 2.3: Histograms of predicted probabilities for IRF vs. SNF from three multivariable logistic regression models that predicted IRF or SNF discharge for acute Medicare stroke patient

General and specific hospital contextual effects

We quantified the magnitude of the overall hospital general contextual effect size using the ICC.[62,78] In the unadjusted model, the ICC was 0.27, which indicated that 27% of the variation in IRF and SNF discharge could be attributed to the hospital (Table 2.5).[79] Somewhat surprisingly, accounting for patient case-mix (i.e. including patient fixed effects in model 2) increased hospital-to-hospital variation (ICC=0.33) (Table 2.5). The ICC of 0.33 for model 2 indicates that hospitals accounted for 33% of the variation in IRF (vs. SNF) discharge, after adjusting for patient level fixed effects. Adjustment for hospital level characteristics (e.g., bed size, for-profit status) in model 3 explained just over a quarter (PVC=28%) of the hospital-to-hospital variation in IRF (vs. SNF) discharge, but the overall general hospital effects were still very large (ICC=0.26) (Table 2.5). There was very large between-hospital variance for hospital level characteristics which is indicated by very wide 80% IORs for all hospital level characteristics (Table 2.5). The 80% IOR for the hospital characteristic of having an IRF unit can be interpreted as follows: Although the average effect of a hospital having an IRF unit increased a patient's odds of receiving discharge to an IRF by 2.53 times (95%CI: 2.25, 2.84) (Table 2.4),[62] the 80% IOR around this estimate was 0.36 to 17.99 which indicates very large between hospital variance in the specific contextual effect of having an IRF unit, with some hospitals having a lower odds of discharge to an IRF and others having a substantially higher odds of discharge to an IRF. Of note, the IOR is a binary level comparison for each specific hospital level factor (i.e., it compares the factor of interest being present vs. not being present. Thus, for categorical variables (i.e., hospital ownership or CMS region) the corresponding 80% IORs are not directly comparable to the aORs from Table 2.4 because different reference groups were used. [62,78]

Table 2.5: Estimates for specific and general hospital effects on influencing IRF or SNF discharge for acute Medicare stroke patients

	All Hospitals
Number of hospitals	1,816
Number of patients	135,414
General hospital contextual effects	
ICC	
Unadjusted	0.27 (0.26-0.28)
Model 2	0.33 (0.31-0.35)
Model 3	0.26 (0.25-0.28)
80% Interval odds ratios (specific hospital contextual effects)	
Bed count (per 50)	0.24-11.88
Hospital ownership	
Church	0.15-7.61
Private-not for profit	0.14-7.08
Private-for profit	0.19-9.43
Government	0.11-5.76
Other	0.14-6.85
Medical school affiliation	0.17-8.29
Affiliated IRF unit	0.36-17.99
Urban (vs. rural)	0.24-11.88
CMS region	
1) CT, ME, MA, NH, RI, VT	0.15-7.67
2) NY, NJ	0.20-9.91
3) DE, DC, MD, PA, VA, WV	0.18-8.81
4) AL, FL, GA, KY, MS, NC SC, TN	0.18-9.24
5) IL, IN, MI, MN, OH, WI	0.14-6.85
6) AR, LA, NM, OK, TX	0.45-22.39
7) IA, KS, MO, NE	0.24-11.92
8) CO, MT, ND, SD, UT, WY	0.14-6.95
9) AZ, CA, HI, NV	0.19-9.29
10) AK, ID, OR, WA	0.09-4.74
PVC	0.28

Abbreviations: CMS: Center for Medicare & Medicaid Services, ICC: Intraclass correlation coefficient, PVC: Proportional change in variance, IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility

Heterogeneity of hospital effects on individual predicted probabilities for IRF (vs. SNF) discharge

To assess heterogeneity of hospital effects on influencing IRF or SNF discharge, we first stratified hospitals based on the propensity to discharge patients to an IRF (vs. SNF). Figure 2.4 shows the ranking of hospitals according to their estimated hospital random intercepts along with their 99% CI's. Just over a quarter of hospitals (n=485, 26.7%) had statistically significant negative random intercepts and were considered SNF favoring hospitals. Around half of hospitals (n=870, 47.9%) were classified as typical hospitals because they had statistically insignificant random intercepts and a quarter were classified as IRF favoring hospitals because they had statistically significant positive random intercepts (n=461, 25.4%).

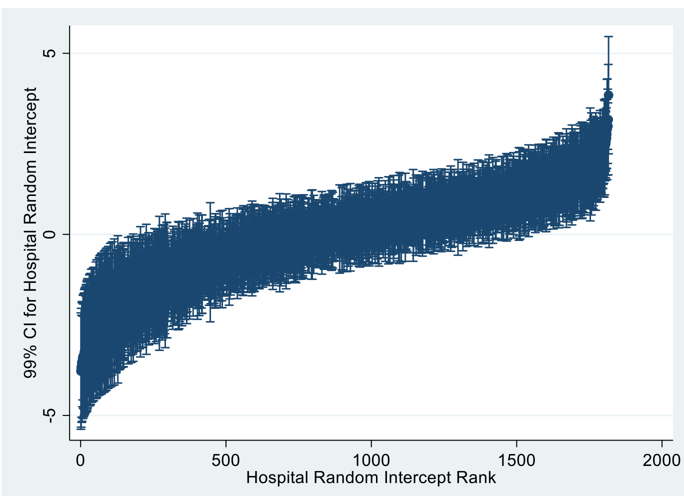


Figure 2.4: Hospital random intercept rank for IRF vs. SNF discharge with 99% confidence intervals (n=1,816 hospitals)

Table 2.6 shows differences in hospital level characteristics between the three types of hospitals. Differences were calculated using one-way ANOVA and chi-square tests and are presented as p-values rather than ASDs because of the three-way comparison. SNF, typical, and IRF favoring hospitals were all very different (i.e., all p-values were <0.01). In general, SNF favoring hospitals were smaller, were more likely to be non-profit, less likely to have an affiliation with a medical school, less likely to be affiliated with an IRF unit, and less likely to be in an urban setting. CMS region 6 (AR, LA, NM, OK, TX) had many more IRF favoring hospitals, while CMS region 10 (AK, ID, OR, WA) had many more SNF favoring hospitals.

Table 2.6: Differences in selected baseline hospital characteristics for hospitals stratified based on their propensity to discharge patients to an IRF or SNF

Hospital characteristics	All Hospitals (n=1,816)	SNF favoring* (n=485, 27%)	Typical¶ (n=870, 48%)	IRF favoring† (n=461, 25%)	p-value
Bed count (SD)	341.3 (247.0)	288.7 (223.3)	329.3 (241.5)	400.5 (261.3)	<0.01
Hospital ownership					<0.01
Church	13.2	10.9	15.2	11.6	
Private-not for profit	43.9	49.1	43.2	41.4	
Private-for profit	16.5	11.1	15.2	22.7	
Government	6.6	7.9	6.7	5.4	
Other	19.9	21.0	19.9	18.9	
Medical school affiliation	42.8	34.2	44.4	46.4	<0.01
Affiliated IRF unit	50.0	22.0	50.1	70.4	<0.01
Urban (vs. rural)	84.1	75.2	84.7	89.7	<0.01
CMS region					<0.01
1) CT, ME, MA, NH, RI, VT	5.8	8.4	6.0	3.6	
2) NY, NJ	8.7	10.4	8.2	8.2	
3) DE, DC, MD, PA, VA, WV	10.3	10.9	10.4	9.6	
4) AL, FL, GA, KY, MS, NC, SC, TN	21.8	25.3	22.8	17.6	
5) IL, IN, MI, MN, OH, WI	18.8	22.0	20.4	13.9	
6) AR, LA, NM, OK, TX	11.3	2.0	7.5	24.7	
7) IA, KS, MO, NE	5.2	2.3	5.8	6.4	
8) CO, MT, ND, SD, UT, WY	2.9	2.0	3.9	1.7	
9) AZ, CA, HI, NV	11.6	10.1	11.6	12.7	
10) AK, ID, OR, WA	3.6	6.6	3.4	1.7	

Abbreviations: IRF: Inpatient rehabilitation facility, SNF: Skilled nursing facility, SD: Standard deviation, CMS: Centers for Medicare and Medicaid Services

*SNF favoring hospitals: Had statistically significant negative random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

¶Typical hospitals: Had statistically non-significant random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

†IRF favoring hospitals: Had statistically significant positive random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

Patient level differences between the three hospital types is shown in Table 2.7. Some pertinent clinically important and statistically significant differences for patients who received care at SNF favoring hospitals were more likely to: be white, have used a SNF in the year prior to their indexed stroke, have dementia, and to be admitted via the ED.

Table 2.7: Differences in selected baseline patient characteristics for hospitals stratified on their propensity to discharge patients to an IRF or SNF

	SNF favoring hospitals* (n=485)	Typical hospitals¶ (n=870)	IRF favoring hospitals† (n=461)	p-value
Number of patients	32,108 (%)	60,987 (%)	42,320 (%)	
Age	82.0 (8.0)	81.4 (8.0)	81.1 (8.1)	<0.001
Race				<0.001
White	82.6	81.5	78.3	
Black	10.3	11.4	12.3	
Hispanic	3.7	3.7	6.3	
Other	3.4	3.4	3.1	
Female sex	61.5	61.0	60.9	0.17
Pre-stroke functional proxies:				
Pre-home-time	358.2 (21.1)	358.4 (21.5)	358.8 (20.4)	<0.001
Previous number of hospitalizations	0.3 (0.7)	0.3 (0.7)	0.3 (0.7)	0.12
Previous SNF use	13.3	11.4	9.7	<0.001
Previous IRF use	0.9	2.5	4.4	<0.001
Comorbidities:				
Total Elixhauser comorbidity index score	4.0 (1.9)	4.0 (1.8)	4.0 (1.8)	0.17
Dementia	9.5	9.2	8.5	<0.001
Stroke characteristics:				
Stroke subtype				0.53
Ischemic	91.0	91.0	90.8	
Intracerebral hemorrhagic	9.0	9.0	9.2	
Stroke severity				<0.001
Mild	39.8	39.0	38.6	
Moderate	38.9	39.6	39.0	
Severe	21.3	21.4	22.4	

Table 2.7 (cont'd)

Hospital Health Services Use				
LOS in days (SD)	5.2 (2.7)	5.1 (2.7)	5.1 (2.8)	0.25
ICU days (SD)	1.7 (2.6)	1.8 (2.6)	1.9 (2.6)	<0.001
CCU days (SD)	0.6 (1.6)	0.6 (1.7)	0.7 (1.8)	<0.001
Emergency department admission	92.7	89.9	89.3	<.001
Lifesaving procedures				
Hemodialysis	1.2	1.2	1.6	<0.001
Gastrostomy tube	5.4	6.2	6.4	<0.001
CPR	<0.1	0.1	<0.1	0.053
Parenteral nutrition	3.0	3.3	2.7	<0.001
Intubation/ventilation	1.6	1.8	1.9	0.001
tPA	5.9	6.6	6.7	<0.001

Abbreviations: LOS: Length of stay, IRF: Inpatient rehabilitation facility, SNF: Skilled nursing facility, ICU: Intensive care unit, CCU: Coronary care unit, SD: Standard deviation

*SNF favoring hospitals: Had statistically significant negative random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

¶Typical hospitals: Had statistically non-significant random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

†IRF favoring hospitals: Had statistically significant positive random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

Table 2.8 shows the general and specific hospital effects for each type of hospital. Overall, classifying hospitals based on their propensity to discharge patients to an IRF or SNF resulted in much smaller general hospital contextual effects for all hospital types which is reflected by much smaller ICCs (ICC range 0.14-0.04) in the fully adjusted models. Notably, the general contextual hospital effects were largest for the SNF favoring hospitals, ICC range 0.20-0.14 for the unadjusted and fully adjusted models (Table 2.8). However, the general hospital contextual effects were almost eliminated for typical and IRF favoring hospitals as the ICCs were all ≤0.05. Compared to the whole sample, the 80% IORs were much narrower which indicates that there is more uniformity in the effect that specific hospital contextual factors had on influencing IRF (vs. SNF) discharge for each hospital type However, notably all the 80% IORs cross 1 which indicates that these characteristics are associated with both an increased and a decreased risk of being discharged to an IRF (vs. SNF).[62]

Table 2.8: Estimates for specific and general hospital contextual effects for influencing IRF or SNF discharge among hospitals categorized based on their propensity to discharge patients to an IRF or SNF

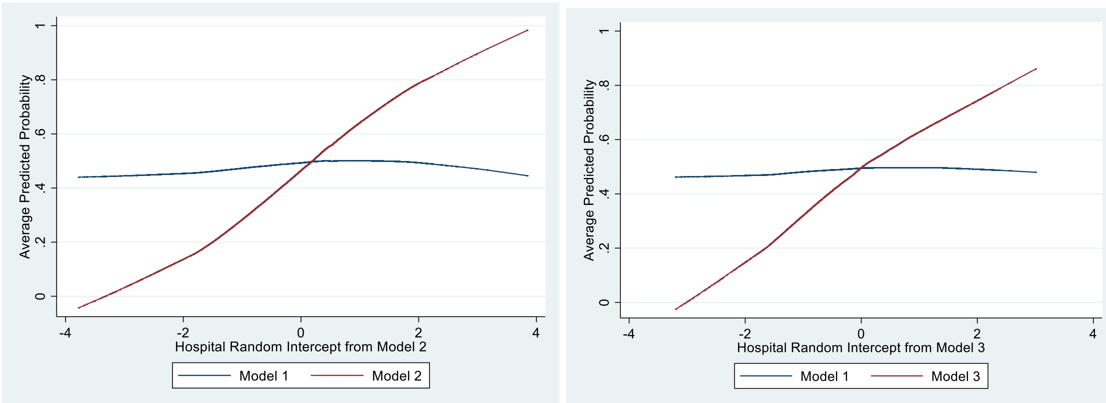
	All Hospitals	SNF favoring hospitals*	Typical hospitals¶	IRF favoring hospitals ǂ
Number of hospitals	1,816	485	870	461
Number of patients (% IRF patients)	135,414 (49.14)	32,108 (22.48)	60,987 (49.39)	43,320 (77.52)
General hospital contextual effects				
ICC				
Unadjusted	0.27 (0.26-0.28)	0.20 (0.18-0.23)	0.05 (0.04-0.06)	0.04 (0.03-0.05)
Model 2	0.33 (0.31-0.35)	0.18 (0.16-0.21)	0.03 (0.02-0.03)	0.05 (0.04-0.06)
Model 3	0.26 (0.25-0.28)	0.14 (0.11-0.16)	0.02 (0.02-0.03)	0.04 (0.03-0.05)
PVC	0.40	0.30	0.13	0.24
80% Interval odds ratios (specific contextual effects)				
Bed count (per 50)	0.24-11.88	0.28-3.82	0.59-1.69	0.51-1.89
Hospital ownership				
Church	0.15-7.61	0.32-4.34	0.58-1.65	0.50-1.85
Private-not for profit	0.14-7.08	0.27-3.67	0.59-1.69	0.52-1.93
Private-for profit	0.19-9.43	0.22-2.94	0.66-1.88	0.58-2.15
Government	0.11-5.76	0.21-2.80	0.59-1.68	0.51-1.89
Other	0.14-6.85	0.28-3.71	0.57-1.61	0.53-1.97
Medical school affiliation	0.17-8.29	0.28-3.71	0.62-1.76	0.52-1.96
Affiliated IRF unit	0.36-17.99	0.38-5.10	0.70-1.98	0.54-2.02
Urban (vs. rural)	0.24-11.88	0.47-6.38	0.59-1.69	0.50-1.85
CMS region				
1) CT, ME, MA, NH, RI, VT	0.15-7.67	0.24-3.20	0.66-1.88	0.54-2.01
2) NY, NJ	0.20-9.91	0.39-5.20	0.66-1.89	0.60-2.23
3) DE, DC, MD, PA, VA, WV	0.18-8.81	0.30-3.98	0.70-2.00	0.46-1.73
4) AL, FL, GA, KY, MS, NC SC, TN	0.18-9.24	0.39-5.21	0.63-1.81	0.46-1.72
5) IL, IN, MI, MN, OH, WI	0.14-6.85	0.27-3.67	0.59-1.69	0.52-1.93
6) AR, LA, NM, OK, TX	0.45-22.39	0.35-4.72	0.79-2.26	0.63-2.35

Table 2.8 (cont'd)

7) IA, KS, MO, NE	0.24-11.92	0.34-4.53	0.67-1.90	0.52-1.95
8) CO, MT, ND, SD, UT, WY	0.14-6.95	0.41-5.54	0.60-1.71	0.54-2.02
9) AZ, CA, HI, NV	0.19-9.29	0.27-3.67	0.65-1.86	0.54-2.03
10) AK, ID, OR, WA	0.09-4.74	0.27-3.66	0.62-1.76	0.38-1.41

Abbreviations: ICC: Intraclass correlation coefficient, PVC: Proportional change in variance, CMS: Centers for Medicare and Medicaid Services, IRF: Inpatient rehabilitation facility, SNF: Skilled nursing facility
*SNF favoring hospitals: Had statistically significant negative random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge
¶Typical hospitals: Had statistically non-significant random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge
†IRF favoring hospitals: Had statistically significant positive random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

Figure 2.5 is a *loess* curve depicting an individual's predicted probability of IRF (vs. SNF) discharge for models 1 and 2 (panel 1) and models 1 and 3 (panel 2) plotted over the ranking of the estimated hospital random intercepts estimated from models 2 and 3 respectively. In each panel, the difference between the estimate from model 1 and from models 2 or 3 can be attributed to the hospital effect on influencing IRF or SNF discharge (panel 1) and further adjustment for hospital characteristics (panel 2).[86] These plots also show that hospitals with negative REs increase the odds of SNF discharge and hospitals with positive REs increase the odds of IRF discharge with large effect sizes the further the hospital RE is away from 0.



Model 1: Patient level F.E.
Model 2: Patient level F.E and hospital R.E
Model 3: Patient, and hospital contextual factors F.E and hospital R.E

Figure 2.5 Average predicted probabilities of IRF (vs. SNF) discharge among Medicare stroke survivors plotted over the hospital random intercepts obtained from the multi-level logistic regression models (n=1816 hospitals)

Table 2.9 quantifies the hospital effect by calculating the proportion of patients that had either a substantial (i.e. >20%), considerable (i.e. >10%), or minimal (i.e. <10%) change in their predicted probabilities ($\hat{p}$) when comparing model 1 to either model 2 or model 3. Any change in $\hat{p}$ between models 1 and 2 can be attributed to the effect of adding the hospital R.E to the multi-level model. Any change in $\hat{p}$ between models 1 and 3 is due to adding both the hospital R.E and hospital fixed effects (i.e., hospital characteristics) to the multi-level model.[86] For the 32,108 patients which were treated at 485 SNF favoring hospitals, adding in the RE for hospital (i.e. $\hat{p}_{\text{model 1}} - \hat{p}_{\text{model 2}}$), led to either a considerable (54.5%) or substantial (29.8%) decrease in their predicted probabilities of IRF discharge. Most (72.6%) of the 60,987 patients at typical hospitals only had minimal changes in their predicted probabilities following the addition of the hospital RE term, but almost all (79%) of the 42,320 patients treated at IRF favoring hospitals had a greater than 10% increase in their predicted probabilities of IRF discharge. Subsequent adjustment for hospital characteristics (i.e. $\hat{p}_{\text{model 1}} - \hat{p}_{\text{model 3}}$), only had a minor effect on the proportion of patients with either substantial or considerable predicted probability changes (Table 2.9).

Table 2.9: Change in the predicted probabilities ($\hat{p}$) of IRF (vs SNF) discharge for hospitals stratified based on their propensity for discharging acute Medicare stroke patients to IRF or SNF

Difference in $\hat{p}_{\text{model 1}}$ - $\hat{p}_{\text{model 2}}$ for SNF favoring, typical, and IRF favoring hospitals						
	> -20 %	-10 to -20%	+/- 10%	+ 10-20%	> +20%	Row total (No. of patients)
SNF favoring hospital (% of row total)	17,602 (54.5%)	9,582 (29.8%)	4,924 (15.3%)	0	0	32,108
Typical hospitals (% of row total)	1,234 (2.0%)	6,893 (11.3%)	44,281 (72.6%)	7,750 (12.7%)	829 (1.4%)	60,987
IRF favoring hospitals (% of row total)	0 (0%)	5 (<0.1%)	8,948 (21.1%)	17,576 (41.5%)	15,791 (37.3%)	42,320
Column total	18,836	16,475	58,153	25,326	16,620	
Difference in $\hat{p}_{\text{model 1}}$ - $\hat{p}_{\text{model 3}}$ for SNF favoring, typical, and IRF favoring hospitals						
	> -20 %	-10 to -20%	+/- 10%	+ 10-20%	> +20%	
SNF favoring hospital (% of row total)	17,638 (54.9%)	9,458 (29.5%)	5,012 (15.6%)	0 (0%)	0 (0%)	32,108
Typical hospitals (% of row total)	1,305 (2.1%)	6,886 (11.3%)	44,146 (72.4%)	7,818 (12.8%)	832 (1.4%)	60,987
IRF favoring hospitals (% of row total)	0	6 (<0.1%)	8,891 (21.0%)	17,579 (41.5%)	15,844 (37.4%)	42,320
Column total	18,943	16,344	49,158	25,397	16,676	

Abbreviations: IRF: Inpatient rehabilitation facility, SNF: Skilled nursing facility
SNF favoring hospitals: Had statistically significant negative random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge
Typical hospitals: Had statistically non-significant random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge
IRF favoring hospitals: Had statistically significant positive random intercepts based on a case-mix adjusted multi-level logistic regression model that predicted IRF (vs. SNF) discharge

DISCUSSION

Overall, we identified that in a large nationally representative database there was significant hospital-to-hospital variation in the proportion of acute stroke patients who were discharged to receive rehabilitation care at an IRF or SNF. At the patient level, several sociodemographic and clinical factors had significant associations with IRF (vs. SNF) discharge. However, consistent with a host of health conditions (e.g., cancer, cardiovascular disease) we identified that the context (i.e., the hospital) of where a patient received care at had a very large effect on influencing the type of rehabilitation care that a patient received.[60,61] In this study, half of all Medicare stroke patients attended an IRF or SNF favoring hospital. Receiving care at one of these hospitals, change the predicted probability of IRF (vs. SNF) discharge by >10% for approximately 80% of patients at these hospitals. We identified several specific hospital factors (e.g. for-profit status, affiliated IRF unit, CMS region) to have large average associations with IRF discharge, but that there was substantial variation in the specific contextual effects of these factors, as reflected by very wide 80% IORs.

Using the conventional quantitative epidemiological approach, our results were consistent with previous analyses.[15–17] At the patient level, we identified several sociodemographic (e.g. age, sex) and clinical factors (e.g. tPA use, higher pre-stroke function) to be associated with IRF (vs. SNF) discharge.[15–17] At the hospital level, we identified that on average, hospitals that were for-profit, had an affiliation with a medical school, had an IRF unit, and were in an urban setting had higher rates of IRF discharge.[15–17] However, the very wide IORs showed that these average effects should be interpreted with caution, because substantial unmeasured heterogeneity likely exists because of the large contextual effects. Several internal and external hospital factors not documented in this study could help explain the very wide IORs. For example, internal hospital factors could include differences in hospital policies (e.g. protocols for rehabilitation assessments), differences in hospitals referral relationships with IRF and SNF units, or differences in clinical culture and clinicians' beliefs on the relative benefits of IRF vs. SNF care.[15,46] External hospital factors could include factors such as regional availability of IRF and SNFs, or the admission policies of specific rehabilitation facilities that hospitals work with.[9]

A notable discrepancy between our patient level findings and those of previous studies was that, apart from dementia we did not identify clinical comorbidities to have significant associations with IRF discharge. This discrepancy was likely the result of only capturing the comorbidities that were coded during the indexed stroke event. In comparison, previous studies used HCC's which capture both inpatient and outpatient claims.[15,17] Thus, we likely undercounted the number of comorbidities that were present for each patient. An important patient level factor that we were unable to account for was proximity of an IRF or SNF to a patient's home.[4] However, a 2019 study by Hong, et al., of 122,084 acute Medicare stroke patients used both HCCs and accounted for IRF and SNF proximity to a patients home reported a patient level adjusted ICC of 0.34.[17] This was very similar to what we found (ICC=0.33) without accounting for those factors. In addition, they also reported the paradoxical finding that case-mix adjustment increased hospital-to-hospital variation and speculated that masking of patient level characteristics was the cause of this observation.[9,17] Alternatively, there are differences in hospital coding intensities for comorbidities and procedures that may have caused case mix adjustment to increase the hospital variability.

A unique contribution of the current study was to explore patient level heterogeneity of hospital effects by stratifying hospitals based on their propensity to discharge patients to an IRF or SNF. Although it is intuitive that SNF favoring hospitals would reduce their probability of patients being discharged to IRFs, we quantified this to show that the hospital effect was to reduce the predicted probability of over 80% of these patients by more than 10%. The result was similar for IRF favoring hospitals, with the exception that these hospitals increased probabilities of IRF discharge.

This study had several limitations which should be considered when interpreting the results. First, although we adjusted for 80 variables, we were unable to adjust for several important unmeasured factors (e.g. patient motivation, social support, patient and family preference, post-stroke function).[19,41] Second, we used administrative data which lacks the granularity of clinical information on medical records and only includes data for Medicare patients.[88,89] However, this study also has several strengths. First, our data source was very large and nationally representative of Medicare stroke patients which provides excellent generalizability. Third, our predicted probabilities were based on an explicitly developed prediction model that used derivation and validation datasets. Finally, we conducted an extensive set of analyses to characterize a broad picture of hospital level variability.

In summary, we used a large nationally representative database to explore heterogeneity of hospital effects on influencing discharge to an IRF (vs. SNF) for acute stroke patients. Overall, we identified several patient and hospital level factors to be associated with IRF discharge, but that these factors were unable to account for the very large general contextual effects that hospitals had on influencing IRF discharge. To understand the impact that these hospital effects have on patient level outcomes future studies should focus on comparative effectiveness of IRF vs. SNF care for acute stroke patients, after accounting for large hospital effects on IRF (vs. SNF) discharge.

CHAPTER 3: SELECTING ACUTE CARE HOSPITALS TO IDENTIFY A TARGET TRIAL POPULATION FOR A PRAGMATIC RANDOMIZED CONTROL TRIAL COMPARING PATIENT OUTCOMES BETWEEN INPATIENT REHABILITATION FACILITIES AND SKILLED NURSING FACILITIES.

BACKGROUND

The RCT is the gold standard research design for clinical evidence. In part because RCTs control for measured and unmeasured confounding biases by using random treatment allocation and thus offer unbiased estimates of treatment effects.[55,90] Historically, most RCTs have been highly explanatory; that is, they are designed to assess treatment efficacy under ideal well controlled circumstances and the magnitude of the treatment effect has been assumed to be constant across treatment centers.[48,91] However, this assumption has been challenged as many of these trials have subsequently been shown to have poor generalizability as site specific characteristics have been shown to influence treatment outcomes.[91] These shortcomings can be addressed in pragmatic RCTs. Pragmatic RCTs maintain random treatment allocation, but aim to maximize generalizability by randomizing a diverse case mix of patients to receive alternative treatments that are delivered in a diverse set of settings that are representative of real world practice.[55,90]

The testing of alternative treatments in representative settings provides real world estimates of effect size which are essential to inform clinical practice and health policy decisions for common medical conditions.[92,30] One such condition is stroke where several large pragmatic trials have recently been conducted.[93–95] With advances in acute care treatment, more attention is being placed on improving the effectiveness of stroke rehabilitation as stroke remains a leading cause of adult disability in the United States.[96] One potential application of a pragmatic RCT in stroke would be to assess the comparative effectiveness of two commonly used alternative rehabilitation settings;[5,6] IRFs and SNFs. IRFs provide intensive rehabilitation over a short period of time (~1-3 weeks), while SNFs provide moderately intensive therapy over a longer period of time (~3-5 weeks).[9,10] There is substantial hospital and geographic variation in the use of these two types of rehabilitation. This large variation has substantial financial implications because IRF care costs approximately double SNF care for stroke patients.[10] Because no RCT has to date compared IRF and SNF outcomes, the existing comparative effectiveness evidence is limited to a handful of observational studies. These studies generally found that patients who were discharged to IRFs had lower mortality rates and increased odds of being discharged home, but these studies have substantial limitations including inconsistent adjustments for known confounders and an inability to adjust for unmeasured confounders.[10,28,33,37,97]

To develop any RCT, substantial planning is required to ensure that a feasible and efficient trial is designed with maximum odds of success.[98] One of the key considerations is the selection of clinical centers or settings.[48,56,90] Settings may be chosen for purely practical reasons (e.g., hospital size, willingness of administrators to participate, availability of resources). A previous qualitative study of 70 trialists considered practical considerations for setting selection to be both common and desirable.[91] Alternatively, settings selection may be aimed at addressing either pragmatic (e.g., obtaining a heterogenous mix of patients and settings) or explanatory (e.g., the ability to collect high quality data, and ability to follow up with patients) components of a trial.[90] Recently, the PRECIS II tools were developed to help trialists optimize the relative pragmatic-explanatory balance to answer the causal question of most interest to patients, clinicians, and stake holders.[48,56]

For a trial comparing IRF vs. SNF rehabilitation for stroke patients there are three important factors to consider when selecting hospitals for a trial. First, hospital referral patterns must be considered because there is substantial variation in the proportion of patients discharged to an IRF (vs. SNF) across hospitals. A typical trial would require 1:1 patient randomization to an IRF or a SNF and hospitals with atypical referral patterns (i.e., hospitals that disproportionality favor discharge to SNFs or IRFs) would be unlikely to participate in the trial as doing so would result in a large deviation from their usual practice. Second, hospital case volume is also critical to ensure efficient patient recruitment. Finally, careful consideration is needed regarding the relative utilization of specific rehabilitation facilities by hospitals. Specifically, selecting hospitals that discharge the majority of their stroke patients to a few larger IRF and SNF facilities would help increase the efficiency the trial as fewer IRF and SNF sites would need to be enrolled in the trial. Also, larger rehabilitation facilities are more likely to have the capacity to participate in such a trial, although excluding smaller facilities would reduce the pragmatic components of the trial.

To inform the design of a pragmatic RCT that compares stroke rehabilitation at IRFs versus SNFs, we aimed to identify a target trial patient and hospital population that would afford an optimal pragmatic-explanatory balance. We explored this balance by assessing the effect that a stepwise application of three practical and explanatory focused hospital-level inclusion criteria had on trial generalizability by comparing target trial patients and hospitals to the starting sample of acute Medicare stroke patients.

METHODS

Patient population

The patient population used in this chapter is the same as described in Chapter 2, with the exception that stroke patients from hospitals with fewer than 20 stroke patients discharged to either an IRF or SNF were not retained because we wanted to start with the full Medicare sample. Specifically, we used Medicare standard analytic files from a 4-year period (2011-2014) to generate a retrospective cohort of community dwelling Medicare fee-for-service ischemic stroke or intracerebral hemorrhagic stroke patients with primary ICD-9 diagnosis codes of 431, 433.x1, 434.x1) who were admitted to an acute care hospital in the US between the two-year period: January 1st, 2012 and December 31st, 2013. From the starting sample of 393,926 patients who were treated at 3,069 hospitals, we excluded patients for the following reasons: 1) had an acute LOS > 14 days (n=13,164), 2) had an inpatient stroke (n=221), 3) had an elective admission (n=11,928), 4) had a current diagnosis of metastatic cancer (n=5,746), 5) received care in a U.S territory (n=1,825), 6) were discharged to a setting other than IRF or SNF (n=207,539), 7) had less than 12 months of continuous Medicare enrollment (n=2,164) and 8) were not part of Medicare Fee-for-service (n=5,970). The resulting sample comprised 145,894 patients who were treated at 3,039 hospitals. Figure 3.1 shows the study flow diagram for how the final retrospective cohort was assembled.

Data sources

We used the same data sources that were described in detail in Chapter 2. This included the following Medicare administrative files: IPC, the inpatient and SNF MedPAR files, CPT file, the MBSF, the ACS, the POS, and Medicare's Hospital Compare database. Additional information on each of these data sources is provided in Supplemental Table 2.1 (Chapter 2). We included data from 2011 until 2014 to allow at least 1 year of information on pre-stroke function/health and at least 1 year of follow-up. The IPC file provided information about the LOS for the acute and IRF stay, as well as ICD-9 diagnosis (including the indexed stroke) and procedure codes for the acute hospitalization. The MedPAR file provided highly aggregated information for a single acute hospital or SNF stay. This file was used to obtain highly categorized charge data for specific charge categories such as emergency department costs or pharmaceutical costs from the inpatient stay, as well as the LOS at the SNF. The MBSF provided information on age, race, sex, enrollment reason, zip code, and disability information from social security. The CPT file was used to identify CPT codes for PT, OT, and SLT provided during the acute in-patient stay. The ACS file provided race and sex specific zip code level aggregate data for information on income and educational attainment.[67] The POS file provided information on hospital characteristics and Medicare's Hospital Compare data provided information on hospital quality by providing patient case-mix adjusted measures of hospital processes and outcomes.[68] Combining these files enabled us to capture all claims at both the acute and rehabilitation facility level. Files were linked using Medicare beneficiary identifiers (for patients) or hospital provider number (for hospitals/rehabilitation facilities).

Outcome

Our primary outcome was IRF vs. SNF discharge after hospitalization for acute stroke care. IRF and SNF patients were identified as patients who were discharged directly to an IRF or SNF and/or who subsequently were admitted to an IRF or SNF within 4 days of hospital discharge. Patients discharged to IRFs and SNFs were identified based on hospital discharge codes 62 and 03, respectively.

Covariates

To address Aim 2 of this dissertation, we only included patient-level variables which were identified as either clinically important or statistically significant (from model 2 in Chapter 2). These variables are listed as follows. Demographic covariates included age, sex, and race (white, black Hispanic, other). Zip code level aggregate data included the annual median household income data (<25k, 25-50k, 50-75k, 75-100k, >100k, missing) which was obtained by linking the patient's zip code of residence to race specific income data from the ACS. Measures of prior health care utilization were taken up to one year prior to the indexed stroke event and included; previous hospitalization (yes/no), home-time (i.e., number of days in last year spent at home and not in a hospital, IRF or SNF),[70] previous use of an IRF (yes/no), and previous use of a SNF (yes/no). Clinical information included the Elixhauser Comorbidity Index (which consists of 31 different comorbidities) and any dementia documented during the indexed hospitalization.[71] Available stroke related information collected during the index hospitalization included stroke subtype (ischemic or intracerebral hemorrhagic) and stroke severity (mild, moderate, severe). Stroke severity was categorized using the stroke administrative severity index.[72] This index is comprised of five ICD-9 discharge diagnostic stroke symptoms (i.e. aphasia, coma, dysarthria/dysphagia, hemiplegia/monoplegia, and neglect) and two ICD-9 procedure codes (i.e. enteral or parenteral nutrition and tracheostomy/ventilation) which were weighted based on the strength of their association with 30 day mortality.[72] This index has been shown to be strongly correlated with the NIH Stroke Scale in Medicare patients.[72] In addition, we used several health services measures as proxies for overall stroke severity. These included LOS, ICU use defined as any stay in the intensive care unit or coronary care unit (yes/no), six lifesaving procedures (i.e., hemodialysis, gastrostomy tube, intubation/ventilation, cardiopulmonary resuscitation, enteral or parenteral nutrition, and tPA use). Hospital charge data (in US dollars) included the amount of laboratory (quartiles 1-4) and pharmacy (quartiles 1-4) services used. The presence of any charge (>$0) was used to identify emergency department admission (yes/no), inhalation therapy services (yes/no), magnetic resonance imaging (yes/no), and operating room use (yes/no). The number of CPT revenue codes (042X-044X) were used as a proxy for the number of acute inpatient physical therapy (0, 1-3, 4-7, 8-11, >11), occupational therapy (0, 1-2, 3-5, 6-7, >7), and speech language therapy (0, 1-2,3-5,6-7,>7) billing codes patients received during their inpatient stay.

Hospital level variables included the number of hospital beds (per 50 bed increase), medical school affiliation (yes/no), hospital ownership (church, private-not for profit, private-for profit, government, other), whether the hospital had an IRF unit directly associated with it (yes/no), whether the hospital was classified as urban or rural, and the 10 CMS geographical regions. We used Medicare's Hospital Compare database[69] to generate composite hospital process and outcome scores. The composite hospital process score was generated by first assigning points to each hospital based on the percentage of patients (missing data=0 points, <90%=1 point, 90-94%=2 points, 95-99%=3 points, and 100%=4 points) at each hospital who received eight stroke services (venous thrombosis prophylaxis, anti-thrombotic use, anti-coagulation use for atrial fibrillation/flutter, any anti-thrombotic use, anti-thrombotic use on day two, discharged on a statin, stroke education, stroke rehabilitation assessment). The points from these eight measures were then summed to create the final composite process score (range 0-32). The composite hospital outcome score (better, no different, worse, missing) classified hospitals as better/worse if they had *either* adjusted 30-day all-cause mortality or adjusted 30-day all-cause readmissions scores which were better/worse than the national average. Further details on technical definitions of patient and hospital factors can be found in Supplemental Table 2.2

Identifying referral networks

To understand the connections between hospitals and rehabilitation facilities we created links in the claims data to identify referral networks consisting of multiple hospital-to-rehabilitation facility dyads. An individual referral network was inferred when a patient's hospital discharge setting matched the setting of their subsequent claim within four days of discharge. We then used case volume (i.e. the number of acute stroke patients discharged to either an IRF or SNF) to first identify the number of IRFs and SNFs that hospitals referred patients to. If at least 5 cases were linked (over the 2-year span) then the hospital-and rehabilitation facility were classified as being part of a *regular use referral network*. Alternatively, if at least 10 cases were linked, then the hospital-and rehabilitation facility were classified as being part of a *frequent use referral network*. Referral networks were categorized using total case volume rather than proportions of patients because case volume is more informative for practical trial design decisions. For each hospital, we calculated how many IRFs and SNFs they referred patients to, as well as how many regular use and frequent use IRF and SNF referral networks they were part of. Finally, hospitals with at least one regular use IRF *and* SNF referral network were considered part of a *regular use referral triad* while hospitals with at least one frequently used IRF *and* SNF referral network were considered part of a *frequent use referral triad*.

Identifying typical hospitals

Because there is significant hospital level variation in IRF and SNF referral patterns,[15–17] we sought to identify hospitals with typical referral patterns when discharging acute stroke patients to receive rehabilitation care at either an IRF or a SNF. We believed that these *typical hospitals* would be more likely to participate in any proposed trial as the 1:1 randomization of patients would not result in large disruptions to their usual referral patterns. To identify typical hospitals, we used three approaches with the goal of identifying the subset of hospitals with minimal hospital level variation in their IRF and SNF referral patterns. In the first non-model-based approach (approach 1) we identified hospitals that had an IRF discharge proportion in the range of 0.20-0.80 as typical hospitals. Approaches 2 and 3, used the same multilevel logistic regression model to model the probability of discharge to an IRF (vs SNF). In this model, we entered all available patient- level covariates as fixed effects and a RE term for hospital. Hospital characteristics were not entered into the model because we sought to focus strictly on patient case-mix adjustment. For each hospital, we estimated the random intercept using the empirical mean Bayes estimate. This estimate is based on the logarithm of the odds ratio for referral to an IRF (vs. SNF) compared to an overall random intercept mean of zero.[83] From this model, we then estimated both 99% (approach 2) and 95% (approach 3) CIs using the standard error of the estimated random intercept. Hospitals were classified as outliers if either their 99% or 95% CIs were either entirely above or entirely below the overall hospital random intercept mean of zero. Hospitals with statistically significant negative random intercepts favored referring patients to SNFs, while hospitals with statistically significant positive random intercepts favored referring patients to an IRF.

Model building and assessment

Model performance was assessed using the area under the curve (AUC) and calibration plots. The AUC was calculated using a Receiver-Operating Characteristic (ROC) curve where the true positive fraction (sensitivity) is plotted against the false positive fraction (1-specificity) for different predicted probability thresholds.[62] The AUC estimates model discrimination because it measures the ability of the model to accurately classify patients who were referred to an IRF vs. a SNF. Model calibration was assessed using a calibration plot which compared observed versus predicted outcomes over 10 deciles of predicted risk. Slopes close to 1 indicate good fit.

For each approach we evaluated the proportion of hospital level variation in IRF and SNF referral patterns by calculating the Δ AUC and the ICC. First, to calculate the Δ AUC we first ran single level and multi-level logistic regression models that adjusted for all available patient level covariates. The Δ AUC ($AUC_{multilevel} - AUC_{single\ level}$) provided an estimate of the added predictive value that was provided by the inclusion of the hospital R.E term.[62] Second, we calculated ICCs for the multilevel models. For these models, the ICC estimated the proportion of the total hospital-level variance in IRF (vs. SNF) referral patterns that were present after patient case-mix adjustment. The ICCs were calculated using the equation, $ICC = \sigma^2 / (\sigma^2 + \frac{\pi^2}{3})$. Where σ^2 is the variation of the hospital random intercepts.

Stepwise application of hospital inclusion criteria to identify hospitals and facilities that optimize the design of the subsequent trial

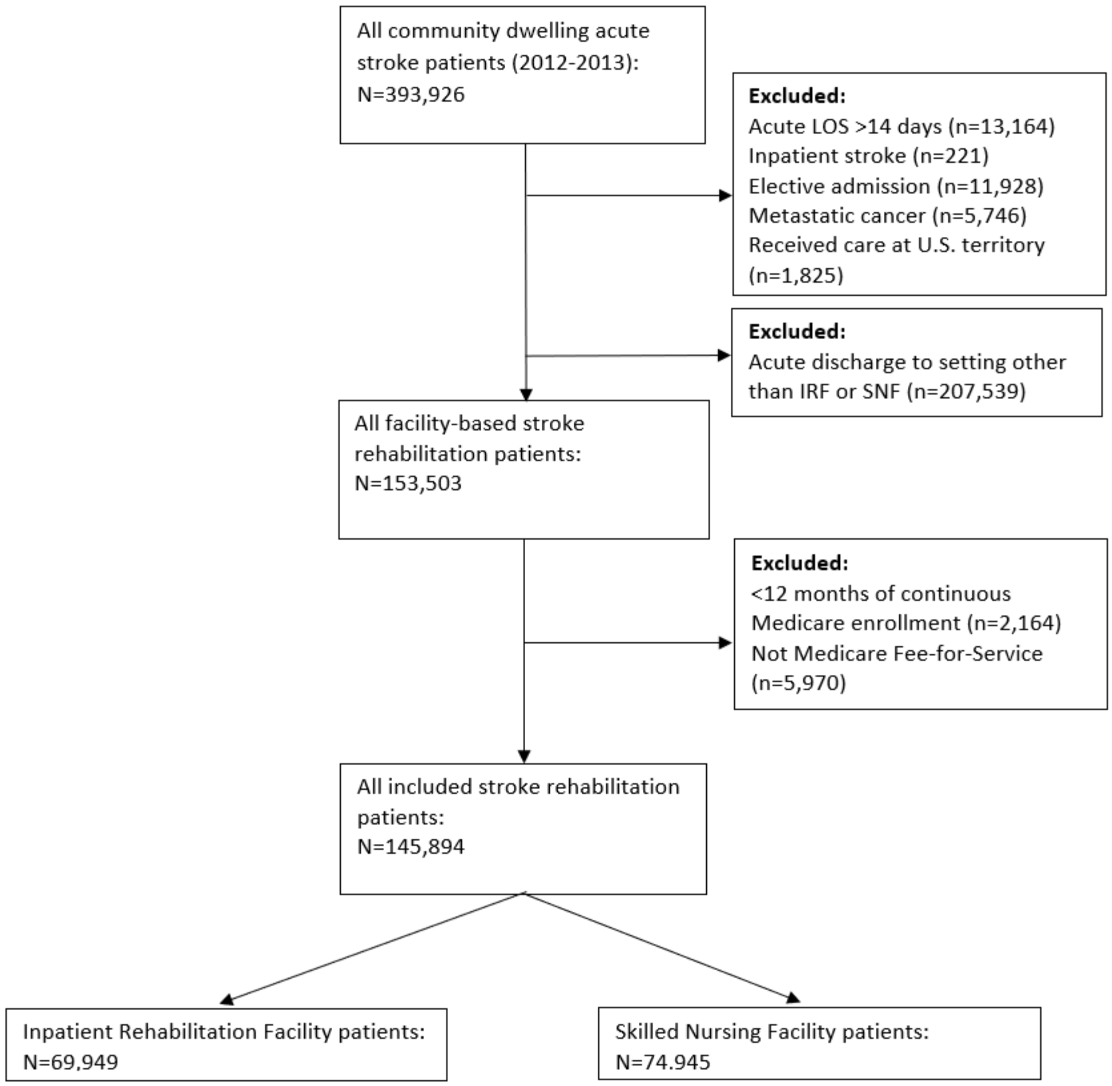
To afford the optimal pragmatic-explanatory balance, we assessed the effects that three types of hospital inclusion criteria had on generalizability. The net effect of applying hospital inclusion criteria affected both patients and rehabilitation facilities as both were nested within a given referral hospital. Generalizability was assessed by calculating the number of hospitals and patients that remained eligible after applying different inclusion criterion. Each inclusion criterion was designed to address practical concerns (i.e., the likelihood of hospitals participating or ease of patient recruitment) and to incrementally increase the explanatory nature (e.g., ability to collect high quality data) of the trial. We used a three-stage approach of applying increasingly restrictive inclusion criteria. First, we only included hospitals with typical IRF and SNF referral patterns (typical hospitals). Typical hospitals had statistically non-significant ($p>0.01$) random intercepts which were estimated from a multi-level logistic regression model that predicted IRF or SNF discharge. Second, we only included typical hospitals with a) more than 20 cases, b) more than 50 cases, and c) more than 100 cases that were referred to either an SNF or an IRF over the 2-year period. Third, we only included typical hospitals which were part of either a *regular or frequent use referral triad* (as previously defined).

Population comparisons

The distribution of patient- and hospital-level factors for both IRF and SNF populations was described using means and standard deviations for continuous variables and percentages for categorical variables. To compare differences in patient- and hospital-level factors between the starting sample and the target trial sample we used ASDs rather than traditional statistically significant testing (p-values) because the former are not affected by the large sample size. We considered ASDs greater than 0.1 to be clinically meaningful.[73] For continuous variables ASDs were calculated using the formula $ASD = \frac{|X_{IRF} - X_{SNF}|}{\sqrt{\frac{S_{IRF}^2 + S_{SNF}^2}{2}}}$. Where $X_{IRF} - X_{SNF}$ is the difference in the sample mean of IRF and SNF patients, and S_{IRF}^2 *and* S_{SNF}^2 are the sample variances for IRF and SNF patients. For categorical variables ASDs were calculated using the formula $ASD = \frac{|p_{IRF} - p_{SNF}|}{\sqrt{\frac{p_{IRF}(1-p_{IRF}) + p_{SNF}(1-p_{SNF})}{2}}}$.[73] where $p_{IRF} - p_{SNF}$ is the difference in the prevalence of the covariate in the IRF and SNF populations respectively.[73]

RESULTS

The initial sample included 145,984 stroke rehabilitation patients which were admitted to 3,039 acute care hospitals and subsequently discharged to 1,150 IRFs and 12,401 SNFs. Details on how the final sample was assembled is shown in the study flow diagram (Figure 3.1).



Abbreviations: LOS: Length of stay
Stroke rehabilitation patients (i.e. cases): Discharged to an Inpatient Rehabilitation Facility or a Skilled Nursing Facility

Figure 3.1: Flow diagram describing the generation of the final study cohort for Aim 2

Patient characteristics for the entire starting population and the ASDs of these characteristics between IRF and SNF patients is shown in Table 3.1. For the starting sample the mean age for patients was 81.5 years, most (80.7%) patients were white, over half (61.2%) were female, and most patients lived in zip codes with race specific median household incomes of less than $75k per annum (Table 3.1). In the year prior to the indexed stroke, 15.8% of patients were hospitalized at least once, 11.4% had used a SNF and 2.7% had used an IRF. Most strokes were ischemic (90.9%) and 21.7% of patients had a severe stroke as classified by the stroke administrative severity data index.[72] During acute hospitalization, the average LOS was 5.1 days and over 90% of patients were admitted through the emergency department. In general, and based off ASD values >0.1, IRF patients tended to be younger, were more likely to be male, treated in the ICU, receive tPA, receive least some OT or SLT (based on receiving >1 CPT code), and receive MRI imaging. Conversely, SNF patients were more likely to have been either hospitalize and/or used a SNF in the year prior to their stroke, had dementia, or received a gastrostomy tube (Table 3.1).

Table 3.1: Differences in baseline patient level characteristics among Medicare stroke survivors discharged to an IRF or SNF

	IRF patients (n=69,949) (%)	SNF patients (n=74,945) (%)	Whole sample (n=145,894) (%)	ASDs*
Sociodemographic characteristics:				
Age (SD)	79.4 (7.7)	83.4 (7.9)	81.5 (8.0)	0.51
Race				
White	80.0	81.5	81.2	0.04
Black	11.6	11.2	11.0	0.02
Hispanic	4.9	4.2	4.3	0.04
Other	3.5	3.1	3.5	0.02
Female sex	56.2	68.5	60.9	0.19
Median annual household income (per $1,000) *				
$<25k	4.0	3.8	3.9	0.02
$25-50k	38.9	39.0	39.0	0.03
$50-75k	36.1	36.4	36.3	0.01
$75-100k	13.1	12.9	12.8	<0.01
$>100 k	6.1	6.2	6.1	0.01
Missing	1.2	1.4	1.9	0.01
Prior health care utilization*				
Previous hospitalization	15.8	24.8	20.5	0.23
SNF use	4.8	17.7	11.4	0.42
IRF use	3.3	2.2	2.7	0.08
Comorbidities:				
Total Elixhauser comorbidity index score (SD)	4.0 (1.7)	4.0 (1.8)	4.0 (1.8)	0.02
Dementia	4.4	13.6	9.21	0.33
Stroke Characteristics				
Stroke subtype				0.01
Ischemic stroke	90.9	91.0	90.9	
Intracerebral hemorrhagic stroke	9.1	9.0	10.1	
Stroke administrative severity index				
Mild	39.6	38.5	39.1	0.02
Moderate	39.2	39.2	39.2	<0.01
Severe	22.2	21.2	21.7	0.03
Hospital Health Services Use				
LOS	5.1 (2.7)	5.2 (2.7)	5.1 (2.7)	0.02
ICU use	60.3	53.5	56.7	0.14
Emergency department admission	89.4	91.4	90.6	0.07

Table 3.1 (cont'd)

Lifesaving procedures				
Hemodialysis	1.1	1.6	1.3	0.04
Gastrostomy tube	3.9	8.2	6.0	0.17
CPR	<0.1	<0.1	0.0	<0.01
Parenteral nutrition	2.4	3.7	2.9	0.07
Intubation/ventilation	1.9	1.7	1.7	0.02
tPA	8.0	5.0	6.1	0.13
Number of physical therapy CPT revenue codes				
0	1.3	1.2	2.3	0.13
1-3	37.7	37.0	37.4	0.02
4-7	37.5	35.1	36.2	0.05
8-11	14.3	15.0	14.7	0.02
>11	9.2	9.5	9.3	0.01
Number of occupational therapy CPT revenue codes				
0	15.1	27.3	21.5	0.30
1-2	31.7	27.7	29.6	0.09
3-6	37.3	30.7	33.9	0.14
7-9	9.5	8.3	8.9	0.04
>9	6.4	6.0	6.2	0.02
Number of speech language therapy CPT revenue codes				
0	21.6	27.1	24.5	0.13
1-2	36.1	32.4	34.2	0.08
3-5	29.2	27.4	28.3	0.04
6-7	6.9	6.9	6.9	<0.01
>7	6.1	6.2	6.2	<0.01
Hospital charge data				
Pharmacy				
Quartile 1	26.3	24.0	25.1	0.05
Quartile 2	25.3	24.6	25.0	0.02
Quartile 3	23.4	26.0	25.0	0.05
Quartile 4	24.5	25.3	24.9	0.02
Laboratory				
Quartile 1	24.9	24.8	24.9	<0.01
Quartile 2	25.8	24.2	25.0	0.04
Quartile 3	25.3	25.0	25.1	0.01
Quartile 4	24.0	25.9	25.0	0.04
Hospital Services use (yes/no)				
Inhalation therapy	35.2	38.7	37.0	0.07
MRI	74.2	64.3	69.0	0.22
Operating room	11.8	12.5	12.2	0.23

Abbreviations: ASD: Absolute standardized difference, SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, CPR: Cardiopulmonary resuscitation, tPA: Tissue plasminogen activator, MRI: Magnetic resonance imaging, CPT: Current Procedural Terminology
Household income estimated from race matched zip code data Prior health care utilization 1 year prior to the indexed stroke. Total Elixhauser comorbidity index: Score range 0-31. ASDs >0.1 were clinically meaningful

Table 3.2 shows selected hospital characteristics for the starting sample which are described at the hospital level. Hospitals had an average of 256 beds, under half (40%) were private-not for profit, just over a third (36.8%) had an affiliation with an IRF unit, a third (33.5) had an affiliation with a Medical school, and most (71.9%) were located in an urban setting. The CMS regions were not equally represented as only 3.2% of hospitals were located in region 10 (AK, ID, OR, WA, while just under a quarter (22.5%) of all hospitals were in CMS region 4 (AL, FL, GA, KY, MS, NC, SC, TN). An equivalent table (with similar distributions) that describes hospital characteristics at the patient level can be found in Supplemental Table 3.1.

Table 3.2: Baseline hospital level characteristics for the 3,039 hospitals that treated and referred 145,894 acute Medicare stroke patients to an IRF or a SNF

	Hospital level (%) (n=3,039)
Number of beds (SD)	256.0 (231.8)
Total hospital process sum score	12.3 (6.7)
Combined mortality and rehospitalizations outcome score	
Worse than national average	3.8
National Average	77.5
Better than national average	3.0
Missing	15.8
Hospital ownership	
Church	10.5
Private not for profit	40.0
Private for profit	18.9
Government	8.2
Other	22.4
IRF affiliated unit	36.8
Medical school affiliation	33.5
Urban hospital	71.9
CMS region	
1) CT, ME, MA, NH, RI, VT	4.5
2) NY, NJ	7.1
3) DE, DC, MD, PA, VA, WV	9.8
4) AL, FL, GA, KY, MS, NC, SC, TN	22.5
5) IL, IN, MI, MN, OH, WI	16.9
6) AR, LA, NM, OK, TX	15.1
7) IA, KS, MO, NE	5.4
8) CO, MT, ND, SD, UT, WY	3.4
9) AZ, CA, HI, NV	11.9
10) AK, ID, OR, WA	3.2

Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, SD: Standard deviation, CMS: Centers for Medicaid and Medicare Services

Total hospital process sum score: Combined score for proportion of patients that received eight stroke quantity process measures

Figure 3.2 is a histogram that shows hospital- level variation of the proportion of patients that each hospital discharged to an IRF or SNF. Overall, around half of the acute stroke patients (48%) were discharged to an IRF. However, there was substantial hospital level variation around this proportion as around 18% of hospitals discharged all their patients to a SNF and around 2%

of hospitals discharged all of their patients to an IRF. An equivalent histogram that presents the distribution of the proportion of patients who were discharged to an IRF which is presented at the patient level (rather than the hospital level) can be found in Supplemental Figure 3.1.

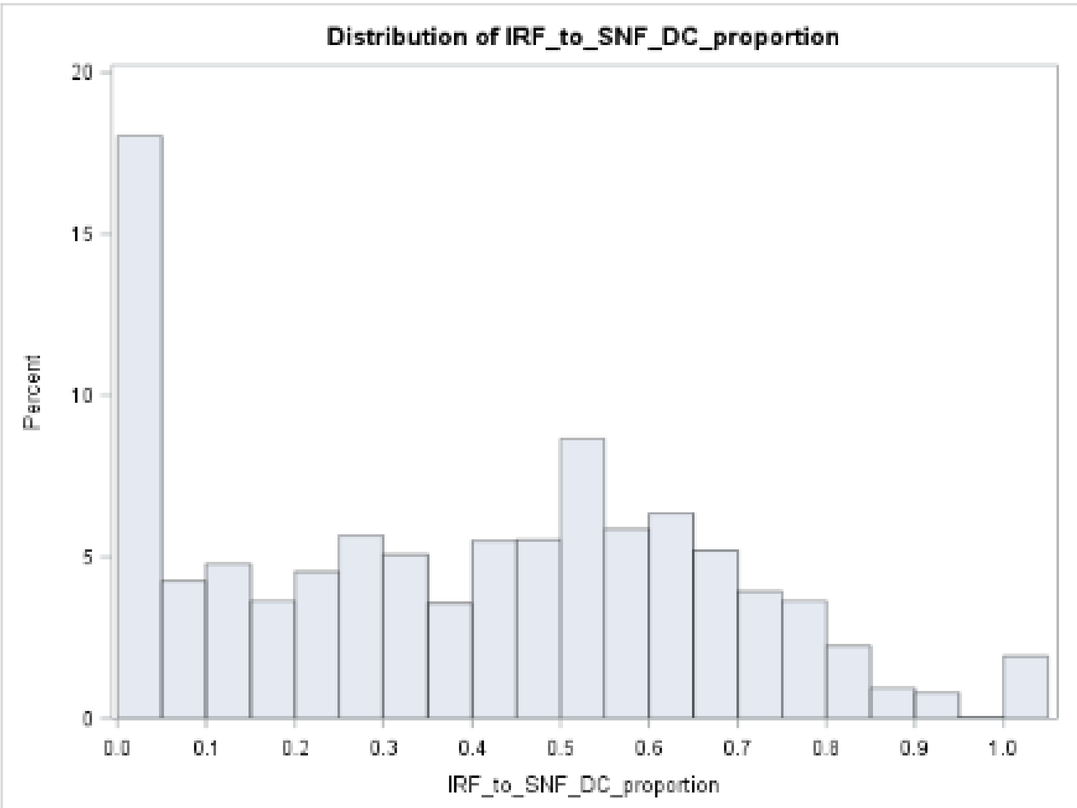


Figure 3.2: Hospital level variation in the proportion of patients (i.e. cases) discharged to an inpatient rehabilitation facility (IRF) compared to a skilled nursing facility (SNF) among the patients who were treated at 3,039 hospitals

Table 3.3 shows information on hospital referral networks for IRF and SNF care. Overall, hospitals discharged at least 1 patient to an average of 2.6 IRF and 9.5 SNF facilities over the 2-year period. On average, hospitals referred at least 5 patients (defined as regular use referral network) to 1.3 IRFs and 2.4 SNFs and hospitals referred at least 10 patients (defined as frequent use referral networks) to an average of 1.2 IRFs and 1.6 SNFs.

Table 3.3: Description of the number of rehabilitation facilities that treated acute stroke patients and hospital referral patterns to these facilities (n=135,415 patients and n=1,816 hospitals)

Rehabilitation Facility and Referral Network Characteristics	
Total number of IRFs	1,150
Total number of SNFs	12,401
Number of IRFs and SNFs hospitals referred patients to	
Mean number of IRFs used by each hospital (SD):	2.60 (2.53)
Mean number of SNFs used by each hospital (SD):	9.50 (9.76)
Number of regular use referral networks	
Mean number of regularly used IRFs used by each hospital (SD):	1.33 (0.82)
Mean number of regularly used SNFs used by each hospital (SD):	2.43 (1.98)
Number of frequent use referral networks	
Mean number of frequently used IRFs used by each hospital (SD):	1.16 (0.50)
Mean number of frequently used SNFs used by each hospital (SD):	1.57 (0.99)

*Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, SD: Standard deviation
Regular use referral network: IRF or SNFs that treated at least five stroke cases that were discharged from a specific hospital
Frequent use referral network: IRF or SNFs that treated at least 10 stroke cases that were discharged from a specific hospital

Stepwise application of hospital inclusion criteria to identify hospitals and rehabilitation facilities eligible for the subsequent trial

Inclusion criteria 1: Hospitals with typical referral patterns

To identify hospitals with typical IRF and SNF referral patterns, we first only included hospitals with more than 20 cases (i.e. acute Medicare stroke patients who were discharged to an IRF or SNF over a 2-year period) because these hospitals were considered large enough to practically serve as candidate trial sites. This criterion left 1,816 hospitals (60.8%) and 135,415 patients (92.8%) to remain eligible for a subsequent trial. Patient and hospital level variation in the proportion of patients that were discharged to an IRF or SNF for this population is shown in Supplemental Figure 3.2. From this sample, using approach 1 (i.e., unadjusted hospital IRF to SNF discharge proportions 0.2-0.8) we identified 1,430 (78.7%) typical hospitals (Table 3.4). The alternate approaches 2 and 3 excluded outlier hospitals with statistically significant random intercepts in the multilevel logistic regression model. Approach 2 (significance set at p<0.01) identified 891 (49.1%) typical hospitals while approach 3 (significance set at p<0.05) identified 665 (36.6%) typical hospitals (Table 3.4). Approach 2 identified more hospitals because setting significance at p<0.01 led to wider CIs, thus making it harder to define a hospital as an outlier.

Table 3.4: Change in the area under the curve (Δ AUC) and intraclass correlation coefficients (ICCs) used to compare the three approaches that were considered to identify hospitals with typical IRF and SNF referral patterns

	AUC from the multilevel logistic regression model	AUC from the single level logistic regression model	Δ AUC	ICC
All hospitals (n=hospital)				
Hospitals with > 20 cases (n=1,816)	0.82	0.72	0.10	0.33
Typical hospitals				
Approach 1 (n=1,430)	0.79	0.72	0.07	0.15
Approach 2 (n=891)	0.77	0.75	0.02	0.04
Approach 2 (n=665)	0.76	0.76	<0.01	0.01

Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF)
All hospitals with > 20 cases (n=1,816 hospitals and 135,415 patients)
Multilevel logistic regression model: Adjusted for all available patient level factors as fixed effects and hospitals as random effects
Approach1 (non-model): Typical hospitals had discharge proportion of 0.2-0.8 (n=1430 hospitals and 116,321 patients)
Approach 2 (statistical model): Hospitals with statistically insignificant random intercepts >0.01 based on the hierarchical logistic regression model (n=891 hospitals and 60,529 patients)
Approach 3 (statistical model): Hospitals with statistically insignificant random intercepts >0.05 based on the hierarchical logistic regression model (n=665 hospitals and 45,581 patients)

The multilevel logistic regression model that was used to identify typical hospitals for approaches 2 and 3 had an AUC of 0.82 which indicates excellent discrimination and the model had a calibration slope close to one which indicates good fit (Figure 3.3).[77] Adjusted ORs between patient and hospital level covariates and discharge to IRF (vs. SNF) from the multilevel logistic regression model are presented in Supplemental Table 3.2.

Receiver operator curve (ROC) Calibration plot

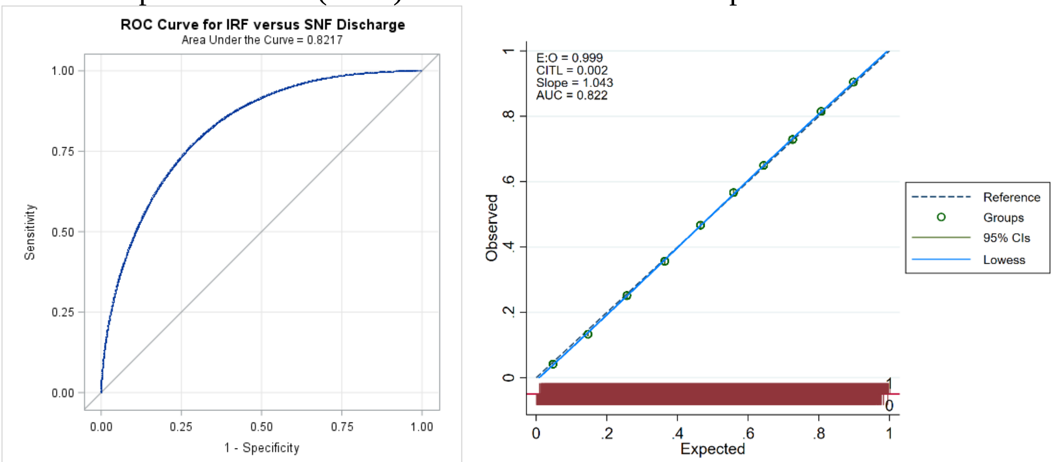


Figure 3.3: ROC and calibration plot from a case mix adjusted multilevel multivariable logistic regression model that predicted inpatient rehabilitation facility or skilled nursing facility discharge for acute Medicare stroke patients (i.e. cases)

Table 3.4 shows the change in AUC (Δ AUC) and ICCs which were the metrics used to compare the three approaches considered in identifying hospitals with typical IRF and SNF referral networks (typical hospitals). Among all 1,816 hospitals with more than 20 cases, the Δ AUC between the single level logistic regression model and multilevel logistic regression model was 0.10. This quantifies the increase in the predictive value that can be attributed to adding the hospital RE term to the multi-level model.[62] Among typical hospitals, the Δ AUC when the hospital RE was added to create the multilevel model was 0.07 for approach 1 (non-model based), 0.02 for approach 2 (statistical significance set at 0.01), and was 0.001 for approach 3 (statistical significance set at 0.05) (Table 3.4). This indicates that there was diminishing predictive value added by knowing which hospital a patient went to (i.e., adding the hospital R.E) for approaches 1 to 3. For all hospitals with more than 20 cases, the ICC was 0.33, which indicates that 33% of the variation in patient referral patterns was attributed to hospital level variation. Approach 1 reduced the amount of hospital variation by half (ICC=0.15). However, approaches 2 and 3 reduced the ICC substantially as hospitals only accounted for 4% (ICC=0.04) and less than 1% (ICC=0.001) for approaches 2 and 3 respectively (Table 3.4). For the remainder of this study, we used approach 2 to classify typical hospitals because these 891 (29.3% of starting sample) hospitals and 60,529 cases (41.5%) offered the best balance between minimal hospital variability in IRF and SNF referral patterns (i.e., ICC=0.004) and maximized the number of eligible patients and hospitals. The remainder of this study focuses on these typical hospitals because we believe that these hospitals would be much more likely to participate in the subsequent trial.

Inclusion criteria 2: Hospital case volume

Table 3.5 shows the number of typical hospitals, patients, and referral patterns by minimal case volumes. For the remainder of the text, we will use the 891 typical hospitals that referred 60,529 stroke cases to 950 IRFs (82.6% of the starting sample) and 7,855 SNFs (63.5% of the starting sample) as the new reference sample. On average, these hospitals referred at least 1 patient each to 2.9 IRFs and 13.5 SNFs. When the criteria were changed to have ≥ 50 stroke cases over 2 years discharged to an IRF or SNF, we retained just over half (n=475, 53%) of the hospitals, but these hospitals treated over three quarters (n=47,326, 78%) of the reference population. However, very few hospitals had case volumes greater than 100 (n=169, 19%) and less than half of patients (n=25,980, 43%) were treated at these hospitals.

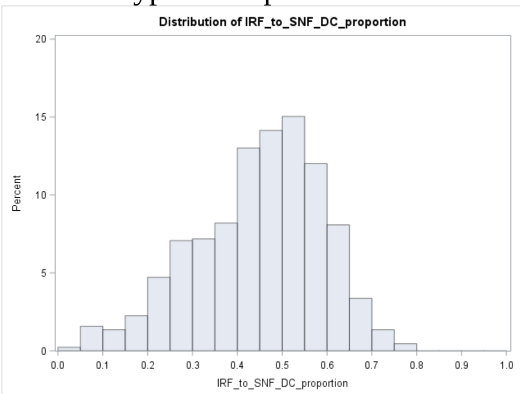
Table 3.5: Number of patients and referral patterns of hospitals with typical IRF and SNF referral patterns over different minimal case volumes thresholds

Minimum case volume (n=hospitals)	≥20 cases (n=891)	≥50 cases (n=475)	≥100 cases (n=169)
Number of patients	60,529	47,326	25,980
% IRF Discharge (SD)	0.47 (0.12)	0.48 (0.10)	0.49 (0.08)
[range]	[0-0.76]	[0.17-0.74]	[0.27-0.64]
IRF referrals			
Number of IRFs that received at least 1 case	950	782	545
Mean number of IRFs used by each hospital (SD):	2.93 (2.57)	3.62 (3.12)	5.20 (4.03)
SNF referrals			
Number of SNFs that received at least 1 case	7,855	6,352	3,932
Mean number of SNFs used by each hospital (SD):	13.51 (10.29)	18.72 (11.45)	28.32 (13.21)

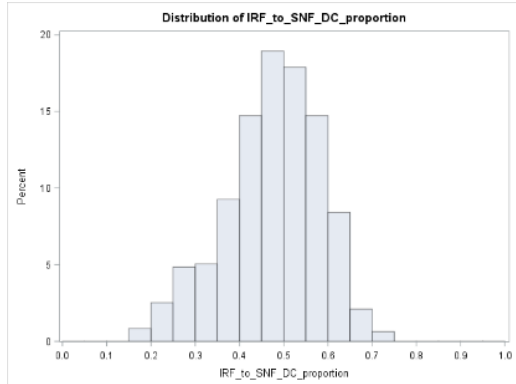
*Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Typical hospitals: statistically insignificant random intercepts based on the multi-level logistic regression model that adjusted for patient level fixed effects and a hospital random effect. *% IRF discharge: Proportion of patients discharged to an IRF versus SNF.

Figure 3.4 depicts histograms that show hospital level variation in the proportion of patients that were discharged to an IRF or SNF among the three types of typical hospitals. The corresponding histograms reported at the patient level are shown in Supplemental Figure 3.3, but the overall pattern of these histograms was similar to those shown in Figure 3.4.

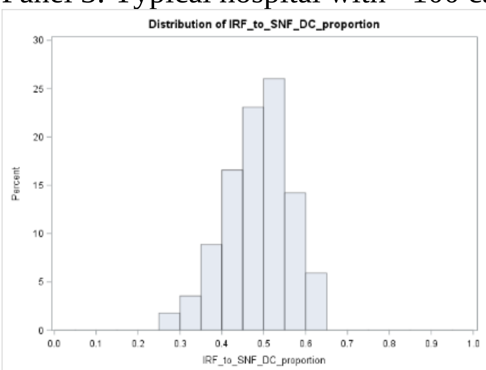
Panel 1: Typical hospitals with >20 cases



Panel 2: Typical hospital with >50 cases



Panel 3: Typical hospital with >100 cases



*case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF)
*Typical hospitals had statistically insignificant (p>0.01) random intercepts from on the hierarchical logistic regression model
Panel 1: 891 hospitals and 60,529 patients
Panel 2: 479 hospitals and 47,326 patients
Panel 3: 169 hospitals and 25,980 patients

Figure 3.4: Hospital-level variation in the proportion of patients (i.e. cases) discharged to an inpatient rehabilitation facility (IRF) compared to a skilled nursing facility (SNF) among patients at typical hospitals

Table 3.6 shows characteristics of *regular use* (i.e., ≥5 patients referred to a specific facility over 2-years) and *frequent use* (i.e., ≥10 patients referred to a specific facility over 2-years) referral networks among typical hospitals by their minimal case volumes. Among the 891 typical hospitals, most (n=823, 92%) were part of at least 1 regular use IRF referral network (mean of 1.29 regular use referral networks). Many hospitals (n=725, 81.3% of typical hospitals) were part of a regular use SNF referral network (hospitals used a mean of 2.55 frequent use SNF referral networks). Of the 7,855 SNFs that received at least one patient, only 1,737 (22%) and

511 (7%) were part of regular use or frequent use referral networks, respectively. Unsurprisingly, as the minimal case volume increases the mean numbers of IRFs and SNFs used also increased. A larger version of this table that also shows the results for the original starting sample of 3,039 hospitals is shown in Supplemental Table 3.3. From this table it can be seen that generally hospitals with larger case volumes have larger referral networks, but the effects vary slightly between all hospitals and typical hospitals.

Table 3.6: Characteristics of regular and frequent used referral networks among typical hospitals by minimal case volume

Minimum case volume (n=hospitals)	≥20 cases (n=891)	≥50 cases (n=475)	≥100 cases (n=169)
Number of patients	60,529	47,326	25,980
Number of IRFs	950	782	545
Number of SNFs	7,855	6,352	3,932
Regular use referral network (≥5 patients referred to a specific rehabilitation facility)			
IRF referral networks			
Number of IRFs	658	480	256
Number of hospitals	823	475	169
Mean number of “regularly used” IRFs by each hospital (SD):	1.29 (0.69)	1.44 (0.85)	1.79 (1.15)
SNF referral networks			
Number of SNFs	1,737	1,338	712
Number of hospitals	725	441	166
Mean number of “regularly used” SNFs by each hospital (SD):	2.55 (1.89)	3.22 (2.10)	4.49 (2.60)
Frequent use referral network (≥10 patients referred to a specific rehabilitation facility)			
IRF referral networks			
Number of IRFs	556	407	197
Number of hospitals	690	460	169
Mean number of “frequently used” IRFs by each hospital (SD):	1.11 (0.41)	1.16 (0.49)	1.35 (0.70)
SNF referral networks			
Number of SNFs	511	424	230
Number of hospitals	354	267	117
Mean number of frequently used SNFs by each hospital (SD):	1.46 (0.84)	1.60 (0.92)	1.97 (1.16)

Table 3.6 (cont’d): Characteristics of regular and frequent used referral networks among typical hospitals by minimal case volume

*Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Typical hospitals: statistically insignificant random intercepts based on the hierarchical logistic regression model

Inclusion criteria 3: Regular use and frequent use referral triads

Table 3.7 shows the effect of only including hospitals that were part of either regular or frequent use referral triads (i.e., defined as hospitals that referred at least 5 or at least 10 patients to both a specific IRF *and* a specific SNF). Among all 891 typical hospitals, three quarters (n=669, 75%) were part of a regular use triad, but less than a third (n=280, 31%) were part of a frequent use triad. Among the 475 typical hospitals with more than 50 cases, most (n=441, 86.5%) of these were part of a regular use triad, and just over half (n=257, 54.1%) were part of a frequent use triad. The larger version of this table that also shows the results for the starting sample is shown in Supplemental Table 3.4

Table 3.7: Number of typical hospitals and patients that are part of regular or frequently use referral triads

Acute care hospitals (n= hospitals)	≥20 cases (n=891)	≥50 cases (n=475)	≥100 cases (n=169)
Number of patients (i.e. cases)	60,529	47,326	25,980
Regular use referral triads (i.e. ≥ 5 patients referred to a specific IRF *and* SNF)			
Number of hospitals	669	441	166
Number of patients (i.e. cases)	52,900	44,950	25,582
Frequent use referral triads (i.e. ≥ 10 patients referred to a specific IRF *and* SNF)			
Number of hospitals	280	257	117
Number of patients (i.e. cases)	29,832	28,890	18,569

Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Typical hospitals had statistically insignificant random intercepts based on the hierarchical logistic regression model.

Final selection of target trial patients and hospitals

Based on the results from Tables 3.5-3.7 my choice of the optimal target trial population included patients who were treated at hospitals with a) typical IRF and SNF referral patterns (i.e., hospitals that had non-significant random intercepts (p>0.01)), b) more than 50 stroke cases discharged over the 2 years, and c) were part of a regular use referral triad. This led to a final target trial population of 441 hospitals (14.5% of the starting sample) and 44,950 patients (30.8%

of the starting sample). These patients were treated at 745 IRFs (64.7% of the starting sample) and 5,974 SNFs (48.2% of the starting sample). In the subsequent chapter we will focus on the effects of only including regular and frequently used rehabilitation facilities. Patient level differences between the starting sample and the target trial population is shown in Table 3.8. Based on ASD values >0.1 at the patient level, the only significant difference was that patients in the target trial population patients (compared to the starting sample) were less likely to have received no (zero) OT CPT therapy codes during their acute hospital stay, but this difference which indicates that the target population was more likely to have received OT during the impatient stay was quite small.

Table 3.8: Differences in patient level characteristics between the starting population and patients identified as being target trial patients

	Starting sample (n=145,984)	Target trial population (n=44,950)	Absolute standardized differences
Sociodemographic			
Age (SD)	81.5 (8.1)	81.5 (8.0)	<0.01
Race			
White	80.7%	82.4%	0.04
Black	11.4%	11.2%	<0.01
Hispanic	4.6%	3.4%	0.06
Other	3.3%	3.0%	0.02
Female sex	61.2%	60.7%	0.01
Median annual household income (per $1,000) *			
< 25k	3.9%	3.7%	0.01
25-50k	39.0%	36.7%	0.05
50-75k	36.3%	37.0%	0.01
75-100k	12.8%	13.5%	0.02
>100k	6.1%	7.3%	0.05
Missing	1.9%	1.8%	0.01
Pre-stroke functional proxies*			
Previous home-time	358.5 (21.2)	358.5 (21.4)	<0.01
Previous number of hospitalizations	0.3 (0.7)	0.3 (0.7)	<0.01
Previous SNF use	11.4%	11.5%	<0.01
Previous IRF use	2.7%	2.3%	0.02

Table 3.8 (cont'd)

Comorbidities			
Total Elixhauser score	4.0 (1.8)	4.0 (1.8)	<0.01
Dementia	9.2%	9.1%	<0.01
Stroke Characteristics			
Stroke subtype			<0.01
Ischemic	91.0%	90.9%	
Intracerebral hemorrhagic	9.0%	9.1%	
Stroke severity*			
Mild	39.1%	38.8%	<0.01
Moderate	39.2%	39.4%	<0.01
Severe	21.7%	21.7%	<0.01
Length of stay	5.1 (2.7)	5.2 (2.7)	<0.01
ICU use	56.7%	57.2%	<0.01
ED admission	90.6%	89.2%	0.05
Lifesaving procedures			
Hemodialysis	1.3%	1.3%	<0.01
Gastrostomy tube	6.0%	6.4%	0.02
CPR	0.0%	0.1%	<0.01
Parenteral nutrition	2.9%	3.6%	0.04
Intubation/ventilation	1.7%	1.9%	0.01
tPA	6.1%	7.1%	0.04
Number of physical therapy CPT revenue codes			
0	2.3%	2.2%	0.01
1-3	37.4%	36.9%	0.01
4-7	36.2%	37.1%	0.02
8-11	14.7%	14.5%	<0.01
>11	9.3%	9.2%	0.01
Number of occupational therapy CPT revenue codes			
0	21.5%	16.9%	0.12
1-2	29.6%	29.9%	0.01
3-6	33.9%	36.4%	0.05
7-9	8.9%	9.8%	0.03
>9	6.2%	7.0%	0.03
Number of speech language therapy CPT revenue codes			
0	24.5%	22.5%	0.05
1-2	34.2%	34.3%	<0.01
3-5	28.3%	29.8%	0.03
6-7	6.9%	7.1%	0.01
>7	6.2%	6.3%	<0.01

Table 3.8 (cont'd)

Hospital charge data			
Pharmacy charge quartiles			
Quartile 1	25.1%	25.8%	0.02
Quartile 2	25.0%	25.3%	0.01
Quartile 3	25.0%	24.8%	0.01
Quartile 4	24.9%	24.1%	0.02
Laboratory charge quartiles			
Quartile 1	24.9%	25.3%	0.01
Quartile 2	25.0%	25.5%	0.01
Quartile 3	25.1%	26.1%	0.02
Quartile 4	25.0%	23.1%	0.05
Hospital Services use (yes/no)			
Inhalation therapy	37.0%	37.4%	<0.01
MRI	69.0%	69.2%	<0.01
Operating room	12.2%	13.6%	0.04

Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator
Starting trial population: All hospitals and patients than met study inclusion/exclusion criteria (n=3,039 hospitals)
Target trial population: Patients who were treated at hospitals that had non-significant random intercepts from a multi-level logistic regression model that predicted discharge to an IRF (vs. SNF), hospitals that had more than 50 stroke rehabilitation patients, and hospitals that discharged at least 5 patients to a specific IRF and a specific SNF (n=441 hospitals)
*Median annual household income taken from race matched zip code data *Pre-stroke functional proxies taken in the year prior to the indexed hospitalization

Table 3.9 shows ASDs for hospital level characteristics between the starting sample of 3,039 hospitals and the 441 target trial hospitals. Based on ASDs >0.1 target trial hospitals were larger, had higher process summary scores, were more likely to be owned by a church or be private not-for profit, and were more likely to be affiliated with a medical school or an IRF unit, and be situated in an urban setting. There were a few important regional differences as target trial hospitals more likely to be located in CMS regions 5 (IL, IN, MI, MN, OH, WI) and 2 (NY, NJ) and less likely to be located in CMS region 6(AR, LA, NM, OK, TX).

Table 3.9: Differences in hospital level characteristics between the starting sample and hospitals identified as being target trial hospitals

	Starting sample (n=3,039)	Target trial population (n=441)	Absolute standardized differences
Number of hospital beds	256.0 (231.8)	429.60 (278.7)	0.73
Hospital stroke process summary score (SD)	12.3 (6.7)	15.9 (4.6)	0.55
Hospital outcome			
Worse than average	3.8	10.0	0.25
No different from average	77.5	80.7	0.08
Better than average	3.0	9.1	0.26
Missing	15.8	0.2	0.60
Hospital ownership			
Church	10.5	19.5	0.25
Private not-for-profit	40.0	49.2	0.19
Private for profit	18.9	8.4	0.31
Government	8.2	4.3	0.16
Other	22.4	18.6	0.09
Medical school affiliation	33.5	50.3	0.35
Affiliated IRF unit	36.8	55.8	0.39
Urban hospital	71.9	93.7	0.60
CMS region			
1) CT, ME, MA, NH, RI, VT	4.5	5.4	0.04
2) NY, NJ	7.1	10.4	0.12
3) DE, DC, MD, PA, VA, WV	9.8	9.5	0.01
4) AL, FL, GA, KY, MS, NC, SC, TN	22.5	24.9	0.06
5) IL, IN, MI, MN, OH, WI	16.9	22.7	0.14
6) AR, LA, NM, OK, TX	15.1	5.4	0.32
7) IA, KS, MO, NE	5.4	4.1	0.06
8) CO, MT, ND, SD, UT, WY	3.4	3.9	0.02
9) AZ, CA, HI, NV	11.9	9.5	0.08
10) AK, ID, OR, WA	3.2	4.1	0.05

Abbreviations: CMS: Centers for Medicare and Medicaid Services
Absolute standardized differences >0.1 were considered clinically important
Starting sample: All hospitals than met study inclusion/exclusion criteria
Target trial hospitals: Hospitals that had non-significant random intercepts from a multi-level logistic regression model that predicted discharge to an IRF (vs. SNF), hospitals that had more than 50 stroke rehabilitation patients, and hospitals that discharged at least 5 patients to a specific IRF and a specific SNF

DISCUSSION

Through the stepwise application of hospital level inclusion criteria, we identified a target trial population for a pragmatic randomized control trial designed to compare the effectiveness of stroke rehabilitation at IRFs compared to SNFs. The final target trial population included 441 hospitals and 44,950 patients, which represent 14.5% and 30.8% of the original starting population. Identification of this population provides important background knowledge to improve trial design efficiency and maximize the odds of trial success.[91] The final selection of hospitals had: 1) typical IRF and SNF referral patterns (based on statistically insignificant hospital random intercepts estimated from a multi-level logistic regression model), 2) more than 50 cases (i.e., acute Medicare stroke patients referred to an IRF or SNF over a 2-year period), and 3) were part of a regular use referral triad (i.e., referred at least five cases to at least one specific IRF *and* at least one specific SNF over a 2-year period). Compared to the starting sample of 3,039 hospitals, the 441 (14.5%) target trial hospitals were not representative of the starting sample (e.g. they had larger case volumes and were more likely to be affiliated with a Medical school and an IRF unit). In contrast, the 44,950 target trial patients were heterogenous with respect to sociodemographic and clinical factors and were very representative of the starting sample of 145,984 patients (there were very few differences as shown in Table 3.8).

Our explicit focus on considering trial generalizability through the selection of specific types of hospitals is somewhat unique and reflects the pragmatic viewpoint of the planned trial. Traditionally, most RCTs have not focused on generalizability issues.[91,99] For example, in a systematic review, Gheorghe, et al., 2013 assessed 129 RCT protocols (with over 300,000 patients) and conducted a qualitative interview of 70 trialists.[91] The authors reported that only 11% of protocols explicitly considered diversity in patient characteristics as a reason for center selection despite many trialists (57%) believing that this was ideal.[91] In another systematic review, Braslow, et al., 2005 included 414 randomized and observational studies and found that only 25% of these studies considered patient representativeness when selecting study centers.[99] Studies characterizing IRF and SNF referral patterns for all medical conditions remain relatively sparse. However, our results for hospital SNF referral patterns were largely similar to a study of 1.5 million SNF referrals for unselected Medicare patients with all types of medical conditions which found that hospitals used a large number of SNFs, but only a few of these SNFs were frequently used.[100]

We chose target trial hospitals with larger case volumes to address practical concerns as larger hospitals would ensure that fewer centers would need to be enrolled. In addition, larger hospitals would be more likely to have the capacity to provide support staff as well as financial and logistical resources to assist in patient recruitment.[101] Generally hospitals with large patient volumes are not representative of hospitals as they have been shown to have superior outcomes across a variety of conditions (e.g., myocardial infarction,[102] surgical procedures,[103] and cancer treatment[104]). However, the relationship between hospital volume and outcomes for stroke patients is mixed. One study of 91,134 acute Medicare stroke patients treated at 625 hospitals participating in the Get With The Guidelines cohort study found no association between bed size or academic status with all-cause mortality or all-cause acute readmissions at either 30 days or 1 year.[105] Another study of 156,886 acute Medicare stroke patients treated at 989 hospitals reported that patients from larger hospitals had worse post-stroke function (measured by the number of days at home as proxy for function).[106] However, these studies were not specific for stroke rehabilitation patients.[9,10]

We included hospitals that were part of regular use IRF *and* SNF referral networks for practical and explanatory reasons. Practically, trial center recruitment is a time and labor-intensive process and it is not feasible to establish contact with an excessively large number of IRFs and SNFs. Within any proposed trial it would be necessary to also enroll specific rehabilitation facilities in addition to the acute hospital. Enrolling these facilities would improve the explanatory nature of the trial because it would be possible to collect more granular process (e.g., therapy type and intensity) and patient (e.g., medical complications, cause of death) level data from each of these rehab facilities. Additionally, establishing a connection with these facilities would likely improve the collection of patient follow-up data.[48,56,90]

Among rehabilitation facilities, the provision of only including regularly used IRFs and SNFs may impact generalizability as only 22% (1,338/5,974) of all the SNFs and 62% (460/745) of all the IRFs from the starting population would be included. To date, most of the literature on rehabilitation facility level variation in quality in case-mix adjusted outcomes has focused on IRFs.[9,23,107] A 2013 study by Graham, et al. of 202,423 stroke patients treated at 717 IRFs identified that about 5% of the variation in home discharge was attributed to facility level factors, with patients treated at larger IRFs faring better.[108] Two large studies (one for IRFs[115] and one for SNFs[116]) of Medicare patients identified that patients who were referred to IRFs and SNFs that were more commonly used by hospitals were less likely to be re-hospitalized and were treated at lower costs, but these studies were not specific to stroke patients.[109,110] However, these studies are all outcome based, and the specific mix of therapeutic processes (i.e., the type, quantity, quality, and intensity of care) that constitutes "usual rehabilitation care" remains poorly characterized.[111,112]

Consistent with previous studies we identified large hospital level variations in IRF and SNF referral patterns for acute stroke patients.[15–17] Through our hospital profiling approach, we found that only half of the hospitals that had at least 20 acute stroke patients that were discharged to and IRF or SNF were identified as typical hospitals based on having a statistically insignificant random intercept ($p>0.001$) estimated from the multilevel logistic regression model. However, there are important practical and ethical reasons why we chose to only include hospitals with typical referral patterns. Ethically, random treatment allocation is predicated on patients being in clinical equipoise.[113,114] Clinical equipoise is established when there is genuine uncertainty within the clinical community towards the optimal treatment choice for a given patient.[113] Clinicians' beliefs of the relative effect of IRF vs. SNF care are shaped by prior experience and clinical culture, thus we believe that outlier hospitals (that have either a high or very low rate of IRF discharge) and/or the clinicians who work at these centers would be less likely to agree to randomize patients. Despite equipoise being established, previous RCTs have been prematurely discontinued because individual clinicians refused to enroll patients who they personally believed should receive one type of treatment.[115] For many trials, clinicians are the "gatekeepers" for patient access, and individual clinicians may abandon a trial protocol if they perceive random treatment allocation could harm their patient (i.e., threaten the moral principle of beneficence).[116,117] Interestingly, in a national study, physiatrists demonstrated substantial variation in their clinical decision making of an optimal choice between IRF versus SNF care for hypothetical stroke patients.[46] Practically, inclusion of these outlier hospitals would also result in substantial deviations from their usual practices which favors the discharge to the majority of stroke patients to one setting over the other. Factors guiding how IRF and SNF referral networks are established are complex, but it is known that they are often underpinned by financial relationships with specific facilities and projections of bed space availability.[118] Because of these factors, we believe that fewer hospital administrators at outlier hospitals would agree to participate in the trial or that there would be less adherence to the research protocol at these hospitals.[56,101]

The current study had several notable strengths. First, our data is highly generalizable as we used a very large national database of all Medicare acute stroke patients. Second, we identified typical hospitals using a multilevel logistic regression model which was effective at reducing hospital level variation compared to unadjusted discharge proportions. Third, our ability to link hospital discharge with IRF and SNF admission provides granular detail on patient flow between these settings and identifies that large established referral networks would be essential for the success of a trial. However, this study also had several notable limitations which should be considered. First, we only included Medicare acute stroke patients which provides a selective view of hospital referral patterns as patients with other insurers may have different referral patterns. Second, we were unable to address clinician and hospital administrator attitudes and beliefs towards perceived trial participation barriers (e.g., questions of additive value of the trial, hospital leadership changes, prioritization of post-acute outcomes, and provider acceptance of the trial).[101] Third, we were unable to adjust for factors which may influence patient participation (e.g., the relative distance of IRF and SNF facilities from a patients home).[15,22]

To conclude, through the stepwise application of hospital and rehabilitation facility level inclusion criteria we identified the target trial population for a pragmatic randomized control trial that compares the effectiveness of stroke rehabilitation at IRFs versus SNFs. We believed that the target trial population afforded the optimal balance between the trial's explanatory and pragmatic components. This target trial population was heterogenous but highly representative of the starting patient population (i.e., acute Medicare stroke patients discharged to an IRF or SNF) but target trial hospitals were not representative at the hospital level as selected hospitals were for example larger, more likely to be affiliated with a medical school, and situated in an urban setting. In the following chapter, we will conduct three comparative effectiveness of IRF vs SNF care starting with this target trial population. The three trials will differ as trial 1 will include all IRFS and SNFs, trial 2 will include regularly used rehabilitation facilities, and trial 3 will include frequently used rehabilitation facilities. We will use the effect size estimates from these trials for sample size calculations

CHAPTER 4: EMULATING A PRAGMATIC CLINICAL TRIAL TO COMPARE THE EFFECTIVENESS OF STROKE REHABILITATION AT INPATIENT REHABILITATION FACILITIES COMPARED TO SKILLED NURSING FACILITIES

BACKGROUND

In clinical medicine, the RCT remains the gold standard of comparative effectiveness research because treatment selection bias is controlled for by random treatment allocation. Ideally, medical guidelines and healthcare policy would be informed by large RCTs, however this is not always practical because RCTs are expensive, time consuming, and there may be ethical considerations associated with randomization.[30] Observational studies are often used to fill these knowledge gaps, but observational studies are limited by the challenges of appropriate statistical adjustment for known confounders and the inability to adjust for unmeasured confounders.[8,28,59]

To address some of the methodological issues with observational data analysis and to help bridge knowledge gaps in the absence of clinical trial data, trial emulation methods have been developed.[59] Emulated trials are hypothetical RCTs where observational data analysis mimics the design features of a true trial (e.g., explicit time zero and synchronized treatment assignment).[52] Randomization is typically emulated using a propensity score. Then, by linking the analysis to the actual idealized trial design, the results reflect a "best guess" of the actual trial results. Emulated trials are relatively new, but examples have used administrative claims data to inform the optimal timing of colon cancer screening,[52] data from large cohort studies to assess antiretroviral treatment switching strategies,[53] and postmenopausal hormone therapy on coronary heart disease.[54] Direct comparisons between emulated and real RCTs are rare, but a recent study found similar results for an actual RCT compared to an emulated trial for the effects of positive-pressure ventilation on oxygen saturation.[119]

In this study, we will use the emulated trial framework to compare the effectiveness of stroke rehabilitation at IRFs versus SNFs for acute stroke patients. Understanding the comparative effectiveness of IRF or SNF care is paramount for several reasons. First, stroke is the leading cause of adult disability in the United States and around half of hospitalized stroke patients are discharged to one of these settings.[5,6,96] Second, there is very large and poorly understood hospital and geographic variation in IRF and SNF use. This variation has garnered increased attention on account that IRF care costs are approximately double those of SNF care and post-acute care has been identified as the largest driver of regional variation in Medicare spending.[10,26] Previous observational comparative effectiveness studies of IRF vs. SNF rehabilitation care for acute stroke patients have generally found that patients treated at IRFs have lower mortality, better physiological and activity level function, and a greater chance of being discharged home.[10,37,120–122]

However, these observational studies are prone to biases because appropriate statistical adjustment for this comparison is complicated by the myriad of patient level factors (e.g. age, sex, health service use prior to stroke) that may affect outcomes and the inability to adjust for other important unmeasured confounders (e.g. community resources, patient and practitioner motivation, quality of rehabilitation care and rehabilitation setting preference).[28] Additionally, almost all studies on stroke rehabilitation base their results as the average difference in outcomes for all patients. However, in the United States stroke rehabilitation care is highly fragmented and hospital contextual effects influence the strength and direction of various selection forces.[15] Thus, it is unclear how generalizable these average differences are for the full population. In contrast, we controlled for these contextual effects by conducting this study within a carefully selected subset of patients, hospitals, and rehabilitation facilities that we believe would represent the ideal target trial population. Given known differences in the quality, type, and intensity of rehabilitation care both within and between the two types of rehabilitation settings we conduct three separate analyses designed to emulate pragmatic trials among increasingly restrictive selection of rehabilitation facilities selected on the basis of their acute stroke case load.[8,27,100,108] The emulation of these three trials will provide a range of effect estimates among different types of facilities that will serve to inform the trial design of a subsequent trial.[98]

METHODS

Patient population

For this chapter, the same starting sample was the target trial population which was identified in Chapter 3. This population consisted of 44,950 patients who were treated at 441 acute care hospitals and subsequently discharged to 745 IRFs and 5,974 SNFs. To review, these patients were identified by using Medicare standard analytic files from a 4 year period (2011-2014) to generate a retrospective cohort of community dwelling Medicare fee-for-service ischemic stroke or intracerebral hemorrhagic stroke patients with primary diagnosis codes (ICD-9), Clinical Modifications of 431, 433.x1, 434.x1) who were admitted to an acute care hospital in the US between the two year period: January 1st, 2012 and December 31st, 2013. From the starting sample of 393,926 patients who were treated at 3,069 hospitals. Patient level exclusion included: 1) had an acute LOS > 14 days (n=13,164), 2) had an inpatient stroke (n=221), 3) had an elective admission (n=11,928), 4) had a current diagnosis of metastatic cancer (n=5,746), 5) received care in a U.S territory (n=1,825), 6) were discharged to a setting other than IRF or SNF (n=207,539), 7) had less than 12 months of continuous Medicare enrollment (n=2,164) and 8) were not part of Medicare Fee-for-service (n=5,970). Hospital level exclusions included: 1) Outlier hospitals (defined as hospitals with statistically significant (p<0.01) random intercepts),

2) Hospitals with fewer than 50 acute stroke patients who were discharged to an IRF or SNF, and 3) Hospitals that were not part of a regular use referral triad. Regular use referral triads were hospitals which discharged at least 5 patients to at least one specific IRF *and* at least one specific SNF. The resulting final starting population that comprised of 44,950 patients (33.1% of the starting population) who were treated at 441 acute care hospitals (14.1% of the starting population) and discharged to 745 IRFs (64.7% of the starting population) and 5,974 SNFs (48.2% of the starting population). This was the starting sample for the first of the three trials.

Data sources

In this chapter, we used the same analytic dataset which was previously described in Chapters 2 and 3. Specifically, this included the following Medicare administrative files: IPC, the inpatient and SNF MedPAR files, CPT file, the MBSF, the ACS, the POS, and Medicare's Hospital Compare database. We included data from 2011 until 2014 to allow at least 1 year of information on pre-stroke function/health and at least 1 year of follow-up. The IPC file provided information about the LOS for the acute and IRF stay, as well as ICD-9 diagnosis (including the indexed stroke) and procedure codes for the acute hospitalization. The MedPAR file provided highly aggregated information for a single acute hospital or SNF stay. This file was used to obtain highly categorized charge data for areas such as emergency department costs or pharmaceutical costs from the inpatient stay, as well as the LOS at the SNF. The MBSF provided information on age, race, sex, enrollment reason, zip code, and disability information from social security. The CPT file was used to identify CPT codes for PT, OT, and SLT provided during the acute in-patient stay. The ACS file provided race and sex specific zip code level aggregate data for information on income and educational attainment.[67] The POS file provided information on hospital characteristics and Medicare's Hospital Compare data provided information on hospital quality by providing patient case-mix adjusted measures of hospital processes and outcomes.[68] Combining these files enabled us to capture all claims at both the acute and rehabilitation facility level. Files were linked using Medicare beneficiary identifiers (for patients) or hospital provider number (for hospitals/rehabilitation facilities).

Covariates

In this chapter we included the same patient level covariates as described in Chapters 2 and 3. These variables include: Demographic covariates included age, sex, and race (white, black, Hispanic, and other). Zip code level aggregate data included the annual median household income data (<25k, 25-50k, 50-75k, 75-100k, >100k, missing) which was obtained by linking the patient's zip code of residence to race specific income data from the ACS. Measures of prior health care utilization were taken up to one year prior to the indexed stroke event and included; previous acute care hospitalization yes/no), home-time (i.e., number of days in last year spent at home and not in a hospital, IRF or SNF),[70] previous use of a IRF (yes/no), and previous use of a SNF (yes/no). Clinical information included the Elixhauser Comorbidity Index (which consists of 31 comorbidities) and any dementia documented during the indexed hospitalization.[71] Available stroke related Information collected during the index hospitalization included stroke subtype (ischemic or intracerebral hemorrhagic) and stroke severity (mild, moderate, severe). Stroke severity was categorized using the stroke administrative severity index.[72] This index is comprised of five ICD-9 discharge diagnostic stroke symptoms (i.e. aphasia, coma, dysarthria/dysphagia, hemiplegia/monoplegia, and neglect) and two ICD-9 procedure codes (i.e. enteral or parenteral nutrition and tracheostomy/ventilation) which were weighted based on the strength of their association with 30 day mortality.[72] This index has been shown to be strongly correlated with the NIH Stroke Scale in Medicare patients.[72] In addition, we used several health services measures as proxies for overall stroke severity. These included LOS, ICU and CCU use (yes/no), six lifesaving procedures (i.e., hemodialysis, gastrostomy tube, intubation/ventilation, cardiopulmonary resuscitation, enteral or parenteral nutrition, and tPA use). Hospital charge data (in US dollars) included the amount of laboratory (quartiles 1-4) and pharmacy (quartiles 1-4) services used. The presence of any charge (>$0) was used to identify emergency department admission (yes/no), inhalation therapy services (yes/no), MRI (yes/no), and operating room use (yes/no). The number of CPT revenue codes (042X-044X) were used as a proxy for the amount of acute inpatient PT (0, 1-3, 4-7, 8-11, >11), OT (0, 1-2, 3-5, 6-7, >7), and SLT (0, 1-2,3-5,6-7,>7) patients received during their inpatient stay.

Description of three target trials

We emulated three different trial designs to provide a range of treatment effect estimates for various target populations defined by the rehab facilities that stroke patients were discharged to. The pre-specified protocol used for the three trials is shown in Table 4.1. The three trials had the same common patient- and acute hospital- level eligibility criteria but had different rehabilitation facility level inclusion criteria. Trial 1 used the same starting trial population as defined above which included all rehabilitation facilities (n=745 IRFs and 5,974 SNFs that treated at least one stroke patient from the 441 hospitals. Trial 2 only included facilities that treated more than 5 patients discharged from a specific hospital (n=460 IRFs and 1,338 SNFs). Trial 3 only included facilities that treated more than 10 patients discharged from a specific hospital (n=297 IRFs and 415 SNFs). Figure 4.1 depicts the study flow diagram for how the starting samples for all three trials were assembled.

Table 4.1: Study protocol for three emulated trials that compared stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

Protocol Component	Description
Eligibility Criteria	
Patient level	Patients: All Medicare community dwelling fee-for-service acute stroke patients discharged to an IRF or SNF from 2012-2013 Patient level Exclusions: *Patient:* acute LOS>14 days, inpatient stroke, elective admission, metastatic cancer, received care at US territory, < 12 months of continuous Medicare enrollment
Hospital level	Hospital level exclusions: Outlier hospitals* that discharged < 50 included stroke patients, and was not part of a regular use referral triad¶
Rehabilitation facilities exclusions (only applicable to trials 2 and 3)	Rehabilitation facility level exclusions: **Trial 1:** All rehabilitation facilities that treated at least one included stroke patient **Trial 2:** Rehabilitation facilities that treated fewer than 5 patients discharged from a single hospital **Trial 3:** Rehabilitation facilities that treated fewer than 10 patients discharged from a single hospital
Treatment	IRF versus SNF stroke rehabilitation
Assignment Procedures	Randomization is emulated via 1:1 propensity score matching: Method=Greedy nearest neighbor, caliper=0.1, and match with replacement
Follow-up Period	1 year following discharge from acute hospital care (t_0)
Outcome:	
Primary	1-year successful community discharge† (yes or no)
Secondary:	a) 1-year all-cause mortality b) Time to successful community discharge c) Time to mortality
Causal Contrast	a) Intention to Treat- based on initial discharge setting (IRF vs. SNF)

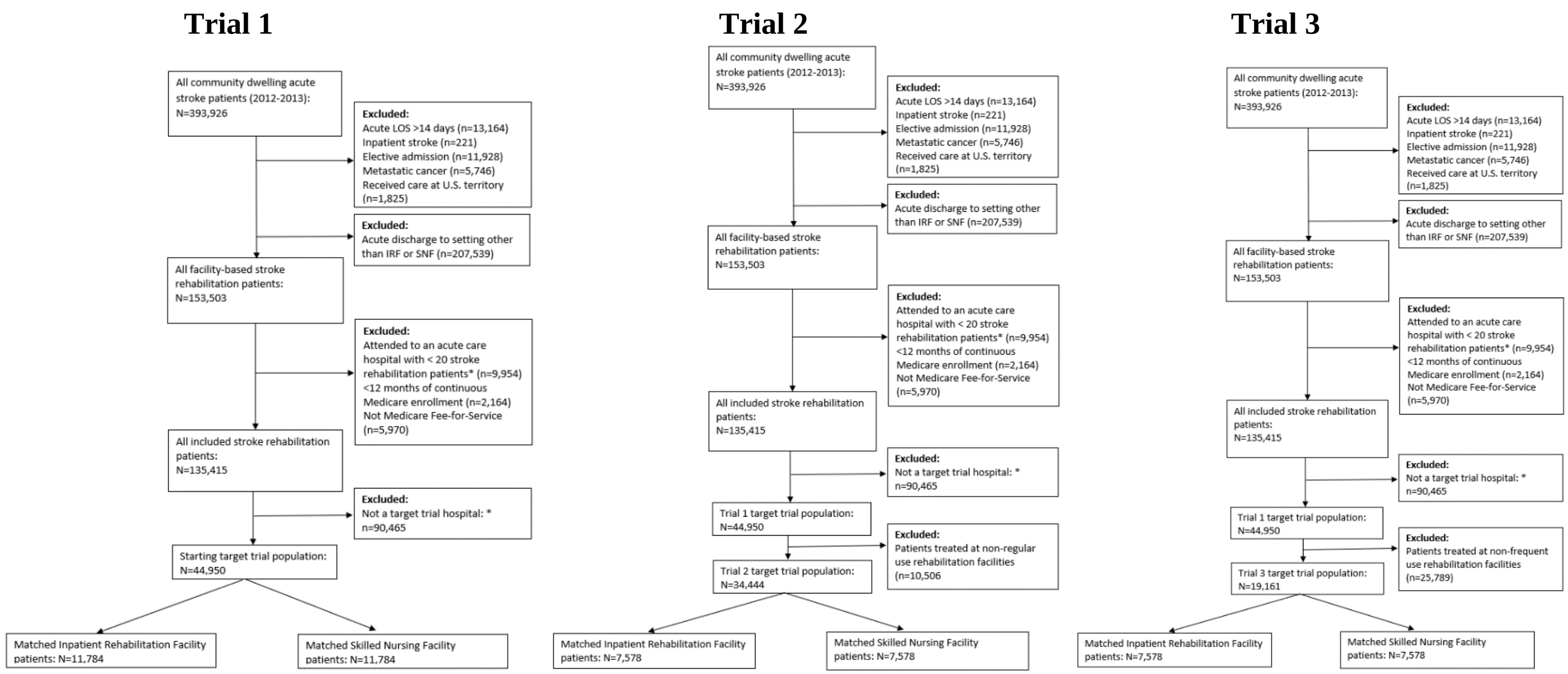
Table 4.1 (cont'd)

Analysis plan	a) Risk difference, relative risks, and odds ratios for binary outcomes b) Kaplan-Meier curves and Cox proportional hazard models for time-to-event outcomes c) Fit a local polynomial regression between the matched pair difference over the propensity score to assess for heterogeneity of treatment effect.[40]
Sensitivity analysis	Competing risks analysis for successful community discharge with death as the competing risk.

*Outlier hospitals: Hospitals with statistically significant positive or negative random intercepts estimates from a multilevel logistic regression model predicting discharge to an IRF or SNF.

¶ Regular use referral triad: Hospitals that discharged at least 5 included patients to a single IRF *and* SNF

† Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF for at least 30 days



*Not a target trial hospial: Hospitals with a) statistically significant positive or negative random intercepts estimates from a multilevel logistic regression model predicting discharge to an IRF or SNF, 2) discharged fewer than 50 Medicare stroke patients to an IRF or SNF over a 2-year period, and 3) Did not discharge at least 5 patients to a specific IRF and SNF over a 2-year period

*Regular use rehabilitation facilitiy (treated at least 5 stroke patients over a 2-year period), frequent use rehabilitaiton facility (treated at least 5 stroke patients over a 2-year period)

Trial 1: All matched patients

Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients

Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.1: Flow diagrams to select participats for three emulated trials that compare stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

Treatment assignment by propensity score matching

Random treatment allocation is the defining feature of a RCT. We emulated randomization by matching IRF and SNF patients with a ratio of 1:1 based on their PS.[49] The PS is the estimated probability of treatment assignment (i.e., IRF or SNF) which was estimated using a single level multivariable logistic regression model that adjusted for all measured baseline characteristics (described in the covariate section above).[51] Matching patients on their PS aims to generate two exchangeable populations (and controls for baseline differences between the populations) as would have been achieved by random treatment allocation.[49,51] We assessed this exchangeability using standardized differences to assess the distribution of baseline covariates between the two groups. We considered any covariate with a standardized difference greater than 0.1 to be poorly balanced.[73]

Prior to matching, the distribution of propensity scores for IRF and SNF patients was assessed, and patients were only matched where common support existed (i.e., the PS distributions overlapped). Because we have large numbers, patients were matched within hospitals on the logit of their propensity score using greedy nearest neighbor matching with a caliper width of 0.1.[123] Matching patients within hospitals helps to control for unmeasured confounders associated with different acute hospitals.[124] Additionally, patients were matched without replacement which resulted in good covariate balance between the IRF and SNF populations in the three samples. However, only around half of IRF patients were matched so we subsequently tried matching with replacement (i.e., more than one SNF patient could be matched to a single IRF patient). Unfortunately, age was poorly balanced between the two groups and despite testing several polynomial terms (e.g., X^2, log(x)) for age, we were unable to improve the balance for age.[125] Thus, we continued with matching 1:1 without replacement as this resulted in well balanced groups for all three trials, albeit with smaller sample size.

Primary outcome

For all outcomes t_0 was considered the point of discharge from the acute care hospital. The primary outcome of interest was the binary outcome of 1-year successful community discharge. This was defined as the proportion of patients discharged home and who remained alive and were not readmitted to an acute hospital, IRF, or SNF for at least 31 continuous days. As part of the 2014 Improving Medicare Post-Acute Care Act, the CMS adapted successful community discharge as a publicly reported quality measure in 2018.[126,107,127] We also analyzed successful community discharge within 90 days instead of 365 days and the time-to-successful community discharge.[107]

Secondary outcomes

Secondary outcomes included all-cause mortality and all-cause acute rehospitalization which were measured at 30 days, 90 days, and 1 year, as well as time-to-death (all-cause mortality).

Primary analysis

All analysis was performed on an intention to treat basis with t_0 set at the date of discharge from the acute care hospital. This ensures that all analyses were conducted using the same timepoint without regard for time spent within each facility or subsequent transitions. The binary outcomes of 1-year successful community discharge and 1-year all-cause mortality differences between treatment arms (IRF vs. SNF) were assessed by calculating the risk difference (RD), risk ratios (RR), and odds ratios (OR) for the unmatched (i.e., all patient eligible for each match) and matched samples. All standard errors used to calculate the 95% CIs were adjusted to account for dependence by using Mantel-Haensel standard error estimates for each matched pair.[51] Binary outcome comparisons for matched pairs were calculated using the "csmatch" and "mcci" commands in STATA version 15.1.[128] Finally, to quantify the potential effects of unmeasured confounding we also calculated E-values for all matched trials. The E-value is the minimum strength of association that an unmeasured confounder (on the relative risk scale) would need to have to explain away the observed treatment effect (after adjustment for all other measured confounders).[129] E-values for relative risk (RR) values greater than 1 were calculated using the formula: $E\ value = \mathrm{RR} + \sqrt{\mathrm{RR}\ \mathrm{x}\ (\mathrm{RR} - 1)}$.[129]

For time-to-successful community discharge and time-to-death, we first used non-parametric estimates to construct either Kaplan-Meier failure (successful community discharge) or Kaplan-Meier survival (all-cause mortality) curves with pointwise 95% CIs.[130] Second, we fit semi-parametric Cox proportional hazard models with treatment arm (i.e. IRF or SNF) as a predictor to calculate crude (i.e. all observations prior to matching) and adjusted (i.e., matched pairs) hazard ratios. The proportional hazard assumption was statistically tested using the Schoenfeld global test. Because of the large sample size, we also visually inspected log-log plots, ensuring the lines were parallel and non-overlapping.[131] We also visually compared observed (Kaplan Meier estimate) vs. predicted (Cox model estimate) survival curves. Close fitting observed and predicted scores indicate closer approximations to the proportionality assumption.[131] For the matched pairs, 95% CIs for the HRs were estimated using bootstrapped estimates of the robust sandwich-type variance estimate to account for clustering between matched pairs. This has been shown to be the least biased method for a matched PS analysis.[132] In addition, to account for the higher death rate among SNF patients, we conducted a competing risks survival analysis. For this analysis, we used a Cox proportional hazard model to semi-parametrically estimate the cause-specific hazard for time-to-successful community discharge (after accounting for the competing risk of death). Standard errors were calculated using a robust sandwich-type variance estimate to account for clustering between matched pairs. Survival analysis was completed using the "sts graph" and "stcox" commands in STATA version 15.1.

Estimating heterogeneity of treatment effect

The result of a matched PS analysis is interpreted as the average treatment effect among the treated (ATT): ATT= E $(Y^1 - Y^0|D=1)$ where E is the expected outcome, Y is the counterfactual outcome for the IRF (Y^1) or SNF (Y^0) populations, and D (1 or 0) corresponds to treatment status with Y=1 for IRF and Y=0 for SNF.[51] This is analogous to the counterfactual framework used by RCTs. However, the average treatment effect among the treated can be misleading when treatment effect size varies systematically across the population.[65,133] Therefore, we evaluated whether heterogeneity of treatment effect was present across the PS by applying a local polynomial regression to the estimated RD of successful community discharge across the PS. The RDs were calculated by the difference in the observed outcome (1 or 0) between each matched pair (IRF patient outcome – SNF patient outcome). We then visually inspected this graph to look for evidence of heterogeneity of treatment effects across different levels of the PS.[133]

Finally, we conducted a limited number of hypothesis generating tests by checking interactions of select baseline covariates with treatment setting in a logistic regression model for the odds of successful community discharge. Dependence was accounted for by using generalized estimating equations with matched pairs treated as clusters.[134] All selected baseline characteristics were identified a-priori as being clinically important (age, race, sex, Elixhauser comorbidity index, stroke subtype, stroke severity, and pre-stroke SNF use) with treatment setting in a logistic regression model for the odds of successful community discharge. All tests were 2-tailed, and significance was set at $p<0.05$.

Sample size calculation

Using the results from all three trials we estimated the anticipated sample size that would be required for a future superiority RCT to detect a statistically significant difference for the primary outcome of 1-year successful community discharge between two independent samples. All tests were 1-tailed (to reflect the hypothesis that IRF care could only have better outcomes than SNF care, not worse), α was set at 0.05 and β was set at either 0.8 or 0.9. Sample size estimates were calculated in STATA 15.1 using the built-in power calculator.

Sensitivity analysis

To explore the effect that unmeasured clinical selection forces which operate across different hospitals (e.g., differences in institutional policies or clinical practice styles) may have had on the estimated treatment effect for IRF vs. SNF care, we conducted a sensitivity analysis in which patients were matched across all hospitals (rather than within each hospital). We conducted this sensitivity analysis using only the patient population from trial 1 and generated the same statistical output from the matched analysis (i.e., RR and RD) with the exception that the survival analysis (hazard ratios (HRs) and survival curves) were not assessed.

RESULTS

As shown in Table 4.2, the mean age of the 44,950 patients included in the starting population (Trial 1) was 81.5 (SD 8.0), the sample was predominantly white (81.2%), and female (60.9%). At the acute hospital the mean LOS was 5.1 (2.7) days and just over half (56.7%) of patients received care in the ICU, and 21.7% suffered a severe stroke as measured by the stroke administrative severity index.[72] The proportion of patients who were discharged to receive either rehabilitation at an IRF (n=21,301, 47.4%) or a SNF (n=23,649, 52.6%) were evenly split. Important differences between the IRF and SNF populations (defined as ASD greater than 0.1) identified that patients treated at an IRF were younger, more likely to be male, and had less pre-stroke healthcare use (e.g., IRF patients were less likely to have been hospitalized and/or used a SNF in the year prior to their stroke. IRF patients were also less likely to have dementia. Generally, patients who received care at an IRF also had lower in hospital health services use including being less likely to receive a gastrostomy tube and were more likely to be in the bottom quartile for total pharmaceutical and laboratory charges. However, the exception for hospital health service use was that that IRF patients were more likely to have received either tPA or an MRI during their acute inpatient stay, and were more likely to have received at least some PT/OT/SLT rehabilitation as measured by the number of CPT revenue codes used (Table 4.2).

Table 4.2: Differences in patient characteristics among Medicare acute stroke patients discharged to receive stroke rehabilitation at either Inpatient Rehabilitation Facilities or Skilled Nursing Facilities

	Whole sample (N=44,950) (%)	SNF patient (N=23,649) (%)	IRF patients (N=21,301) (%)	Absolute standardized differences *
Age	81.5 (8.0)	83.5 (7.8)	79.1 (7.6)	0.57
Race				
White	81.2	82.5	82.3	0.01
Black	11.0	11.4	10.9	0.02
Hispanic	4.3	3.4	3.5	0.01
Other	3.5	2.7	3.3	0.04
Female sex	60.9	65.9	54.9	0.23
Median annual household income (per $1,000) ¶				
$<25k	3.9	3.8	3.7	0.01
$25-50k	39.0	36.8	36.6	<0.01
$50-75k	36.3	37.1	36.9	0.01
$75-100k	12.8	13.6	13.5	<0.01
$>100 k	6.1	7.0	7.6	0.02
Missing	1.9	1.7	1.9	0.01
Prior health care utilization†				
Pre-stroke home-time	358.46 (21.41)	355.1 (27.2)	362.1 (11.0)	0.33
Prior hospitalization	20.4	25.2	15.0	0.26
SNF use	11.4	17.7	4.7	0.42
IRF use	2.7	2.0	2.6	0.04
Comorbidities:				
Total Elixhauser comorbidity score	4.0 (1.8)	4.0 (1.9)	4.0 (1.8)	0.02
Dementia	9.21	13.7	4.1	0.34
Stroke Characteristics				
Stroke subtype				0.02
Ischemic	90.9	91.2	90.7	
Intracerebral hemorrhagic	10.1	8.8	9.3	
Stroke administrative severity index				
Mild	39.1	39.0	38.7	<0.01
Moderate	39.2	39.5	39.3	<0.01
Severe	21.7	21.5	22.0	0.01

Table 4.2 (cont'd)

Hospital health services use				
Length of stay (days)	5.1 (2.7)	5.2 (2.7)	5.1 (2.7)	0.02
ICU use	56.74	55.2	59.3	0.08
Emergency department admission	90.6	89.9	88.4	0.05
Lifesaving procedures				
Hemodialysis	1.3	1.6	0.9	0.06
GI tube	6.0	8.8	3.8	0.21
CPR	0.0	0.1	0.0	0.01
Nutrition	2.9	4.3	2.9	0.07
Intubation/ventilation	1.7	1.8	2.0	0.02
tPA	6.1	5.4	9.0	0.14
Number of physical therapy CPT revenue codes				
0	2.3	3.3	1.1	0.15
1-3	37.4	37.4	36.4	0.03
4-7	36.2	35.7	38.7	0.06
8-11	14.7	14.4	14.7	0.01
>11	9.3	9.2	9.2	<0.01
Number of occupational therapy CPT revenue codes				
0	21.5	22.4	10.7	0.32
1-2	29.6	28.9	31.1	0.04
3-6	33.9	33.1	40.0	0.14
7-9	8.9	9.1	10.6	0.05
>9	6.2	6.5	7.6	0.04
Number of speech language therapy CPT revenue codes				
0	24.5	24.7	20.0	0.11
1-2	34.2	32.8	36.0	0.07
3-5	28.3	29.1	30.6	0.03
6-7	6.9	7.2	7.0	<0.01
>7	6.2	6.2	6.3	<0.01
Hospital charge data				
Pharmacy				
Quartile 1	25.1	23.1	28.8	0.13
Quartile 2	25.0	24.5	26.2	0.04
Quartile 3	25.0	27.0	22.4	0.10
Quartile 4	24.9	25.4	22.6	0.07
Laboratory				
Quartile 1	24.9	21.6	29.5	0.18
Quartile 2	25.0	24.2	27.0	0.06
Quartile 3	25.1	27.3	24.8	0.06
Quartile 4	25.0	27.0	18.7	0.20

Table 4.2 (cont'd)

Hospital Services use (yes/no)				
Inhalation therapy	37.0	38.3	36.4	0.04
MRI	69.0	65.0	74.0	0.20
Operating room	12.2	14.3	12.8	0.05

Abbreviations: IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility, ICU: Intensive care unit, GI: Gastrostomy tube, tPA: Tissue plasminogen activator, MRI: Magnetic Resonance Imaging, CPT: Current procedural terminology

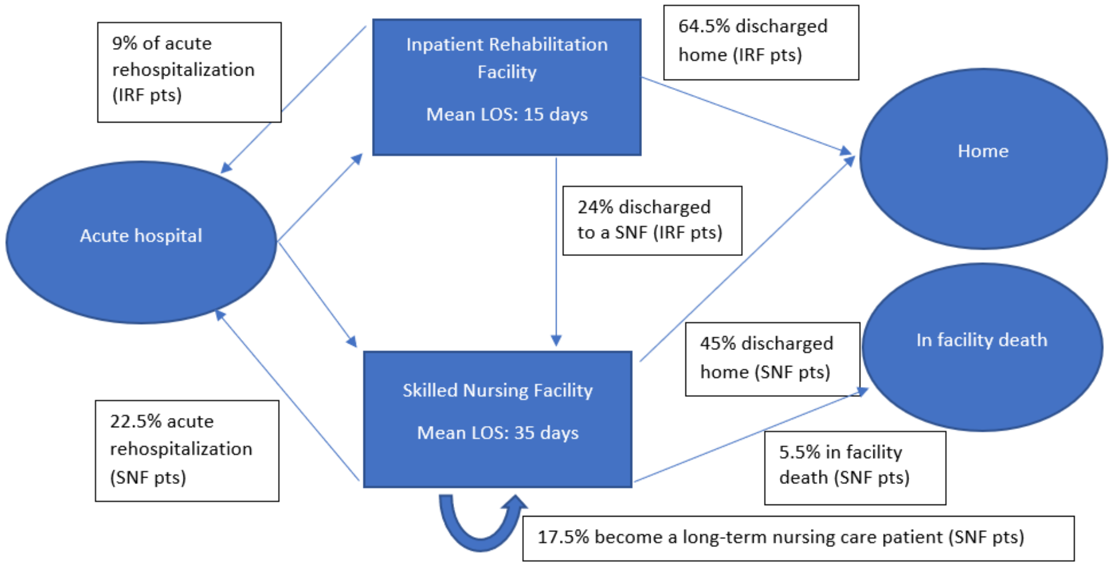
* Absolute standardized differences >0.1 considered clinically important

¶Median annual household income: taken from race matched zip code data Prior health care utilization

† Taken 1 year prior to the indexed stroke event

Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator, CPT: Current Procedural Terminology

Figure 4.2 shows common discharge destinations for the first discharge setting (i.e., IRF or SNF) for the full starting cohort following admission to either an IRF or SNF. Supplemental Table 4.2 shows the mean LOS and interquartile ranges prior to each discharge setting. Among patients treated at an IRF, the average LOS was 15 days [IQR: 9-20] and most (64.5%) patients were subsequently discharged home. Ten percent of IRF patients were discharged back to the acute hospital (9.1%) while a quarter were discharged to a SNF for further rehabilitation care (24.4%). Very few died within the IRF (0.2%). Among patients treated at a SNF, the average LOS was 35 days [IQR: 13-47] and just under half of these patients (45.4%) were subsequently discharged home. Among the remaining patients, 22.5% were readmitted back to the acute hospital, 17.5% transitioned to become a long-term nursing home resident at the same facility, 5.5% died at the SNF, and 3.1% were discharged to another SNF. Differences in the starting samples between the three trials are shown in Supplemental Table 4.1.



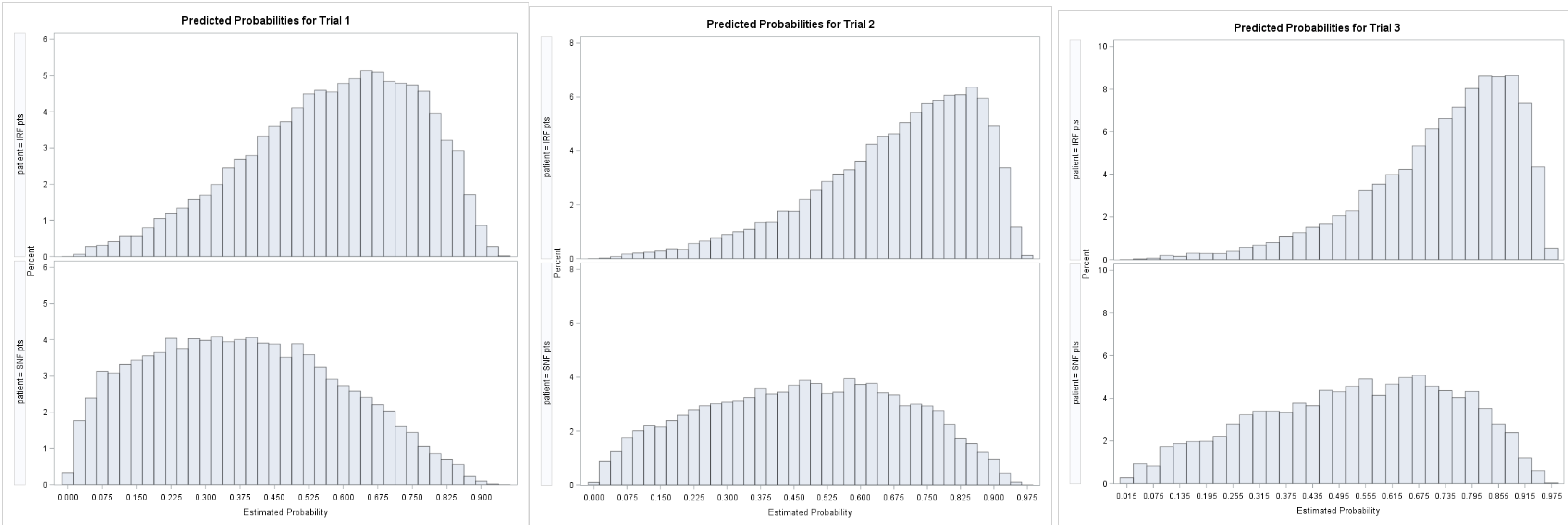
Abbreviations: LOS: Length of stay, pts: patients, IFR: inpatient rehabilitation facility, SNF: skilled nursing facility
Note: 0.2% of IRF patient died infacility an 0.3 % of SNF patient were discharged to a different SNF

Figure 4.2: First patient discharge destination following treatment at the initial rehabilitation facilities (Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF)

Matched samples

Based off the distribution of the PS for IRF and SNF patients (Figure 4.3), the range of common support was defined as a PS between 0.1 and 0.9 – thus patients were matched only in this range. For trial 1, 23,568 patients (11,784 pairs) were matched. These subjects were treated at 662 IRFs and 4,579 SNFs (Supplemental Table 4.3). For trial 2, 15,156 patients (7,578 pairs) were matched and were treated at 442 IRFS and 1,319 SNFs. Finally, for trial 3, 7,456 patients (3,728 pairs) were matched and were treated at 254 IRFs and 414 SNFs. All three trials had similar baseline patient characteristics. Because of the three-way comparison, ASDs were not used and p-values are confounded by the large sample size so differences >1% were considered clinically important.[135] Based on differences >1%, patients in trial 1 were more likely to have dementia and had higher use of previous SNFs, ICUs, gastrostomy tubes, and operating rooms. Further details for the baseline characteristics for each starting trial population as well as the number of hospitals, and rehabilitation facilities included in each trial are shown in Supplemental Tables 4.1 and 4.3. All trials were well balanced between the IRF and SNF treated patients based on the standardized differences for all covariates being <0.1 (Figure 4.4 and Supplemental Table 4.4).

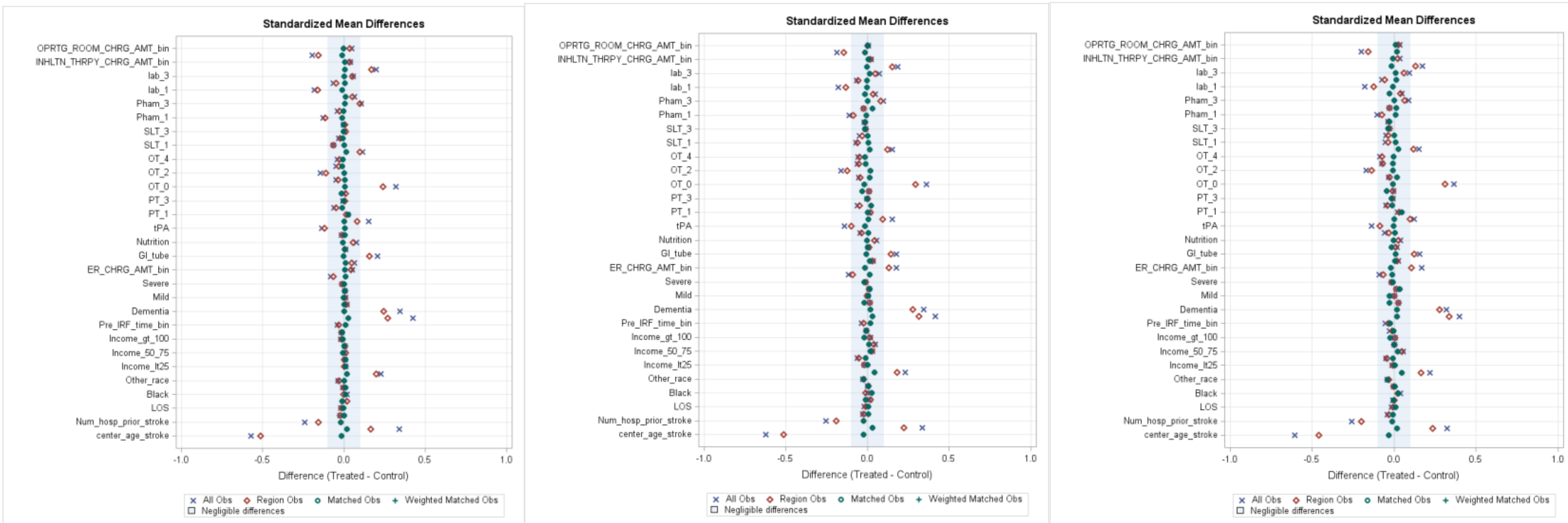
Trial 1 Trial 2 Trial 3



Zone of common support: 0.10-0.90
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.3: Distribution of the probability of discharge to an Inpatient Rehabilitation Facility (versus a Skilled Nursing Facility) estimated from a patient level logistic regression model

Trial 1 Trial 2 Trial 3



Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients
Shaded zone in middle represents zone of clinically irrelevant differences in standardized differences (i.e., >-0.1 and <0.1)

Figure 4.4: Standardized differences of patient level covariates after Inpatient Rehabilitation Facility and Skilled Nursing Facility patients were matched based on their estimated propensity score

Descriptive outcomes

Table 4.3 shows descriptive outcomes for IRF and SNF patients for the three trials. In trial 1, 81.6% and 60.6% of IRF and SNF patients achieved successful community discharge within 1 year. In trials 2 and 3, the event rates for successful community discharge were similar for IRF patients (80.6% and 80.3% respectively) but were higher for trials 2 and 3 for SNF patients (63.4% and 68.0% respectively). The similar event rates for IRF patients and different event rates for SNF patients indicates that any variation in treatment effect across the trials was driven by differences within SNFs. Overall, over three quarters of patients who were eventually successfully discharged to the community did so within 90 days for both IRF and SNF patients. For 1-year all-cause mortality, SNF patients consistently had higher mortality rates (32.5%, 32.5%, and 31.1%) compared to IRF patients (20.7%, 21.8%, and, 22.6%) for trials 1, 2, and 3, respectively. Finally, for all-cause acute readmissions, SNF patients had slightly higher readmissions for all timepoints, but the differences between the two populations was less than a 3% at all timepoints (Table 4.3).

Table 4.3: Descriptive outcomes for the three propensity score matched target trials comparing stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

	Trial 1 (n=11,784 matched pairs)		Trial 2 (n=7,578 matched pairs)		Trial 3 (n=3,728 matched pairs)	
	IRF (%)	SNF (%)	IRF (%)	SNF (%)	IRF (%)	SNF (%)
Successful community discharge						
90 day	68.1	45.2	67.3	48.0	66.6	53.3
1 year	81.6	60.6	80.7	63.4	80.2	68.0
All-cause mortality						
30 day	3.0	9.2	3.1	8.4	3.5	7.1
90 day	8.1	17.2	8.4	16.5	9.4	15.2
1 year	20.7	32.5	21.8	32.5	22.6	31.1
All-cause acute hospital readmissions						
30 day	13.4	16.1	13.8	16.5	14.0	16.4
90 day	25.9	28.9	27.1	29.6	28.4	30.2
1 year	48.6	49.2	49.8	50.8	50.3	51.2

Time starts upon discharge from the acute care hospital
Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Comparative outcomes for binary endpoints

Table 4.4 shows outcome comparisons between the two treatment arms for the three trials. In trial 1, the unadjusted (i.e., all eligible patients) RD for successful community discharge was 0.34 and the RR was 1.82 (95% CI: 1.79, 1.85). A RD of 0.34 indicates that if 100 patients were discharged to an IRF rather than a SNF, then an additional 34 of these same patients would be expected to be successfully discharge home within 1 year. A RR of 1.82 indicates that patients who were treated at an IRF were 82% more likely to be successfully discharged back to the community compared to patients who were treated at a SNF. Adjustment by PS matching reduced these differences substantially, but patients treated at an IRF were still 35% more likely (RR: 1.35, 95%CI: 1.32, 1.37) to be successfully discharged home compared to patients who were treated at a SNF. For the matched trials 2 and 3, the RD (0.17 and 0.12) and RR (1.27 and 1.18) estimates were both lower (especially trial 3) compared to trial 1. For 1-year all-cause mortality, the RD (0.12, 0.11 and 0.08) and RR (0.85, 0.86, 0.89) estimates for the matched trials 1, 2, and 3 showed a significant (albeit much smaller) treatment effect for IRF vs. SNF care. Overall, for both successful community discharge and 1-year all-cause mortality, when examining the results using the RD the net effect of adjustment using PS matching was very similar across the three trials representing a downward shift of about 15% on the absolute scale (Table 4.4).

In addition, we also calculated E-values for the primary outcome of successful community discharged for trials 1, 2, and 3. For trial 1, the E-value was 2.04 which indicates that on the relative risk (RR) scale, an unmeasured confounder of 2.04 would be needed to nullify the emulated trial's observed significant RR effect size of 1.35. When applied to the observed lower bound of the 95% CI for this RR (i.e., lower confidence level= 1.32) the unmeasured confounder would have to be 1.97 to result in a non-significant estimate. The equivalent E-values for the RR estimates for trial 2 (RR= 1.27) and 3 (RR = 1.18) were 1.86 and 1.64, respectively, and the equivalent estimates for the lower bounds of the RR were 1.81 and 1.57.

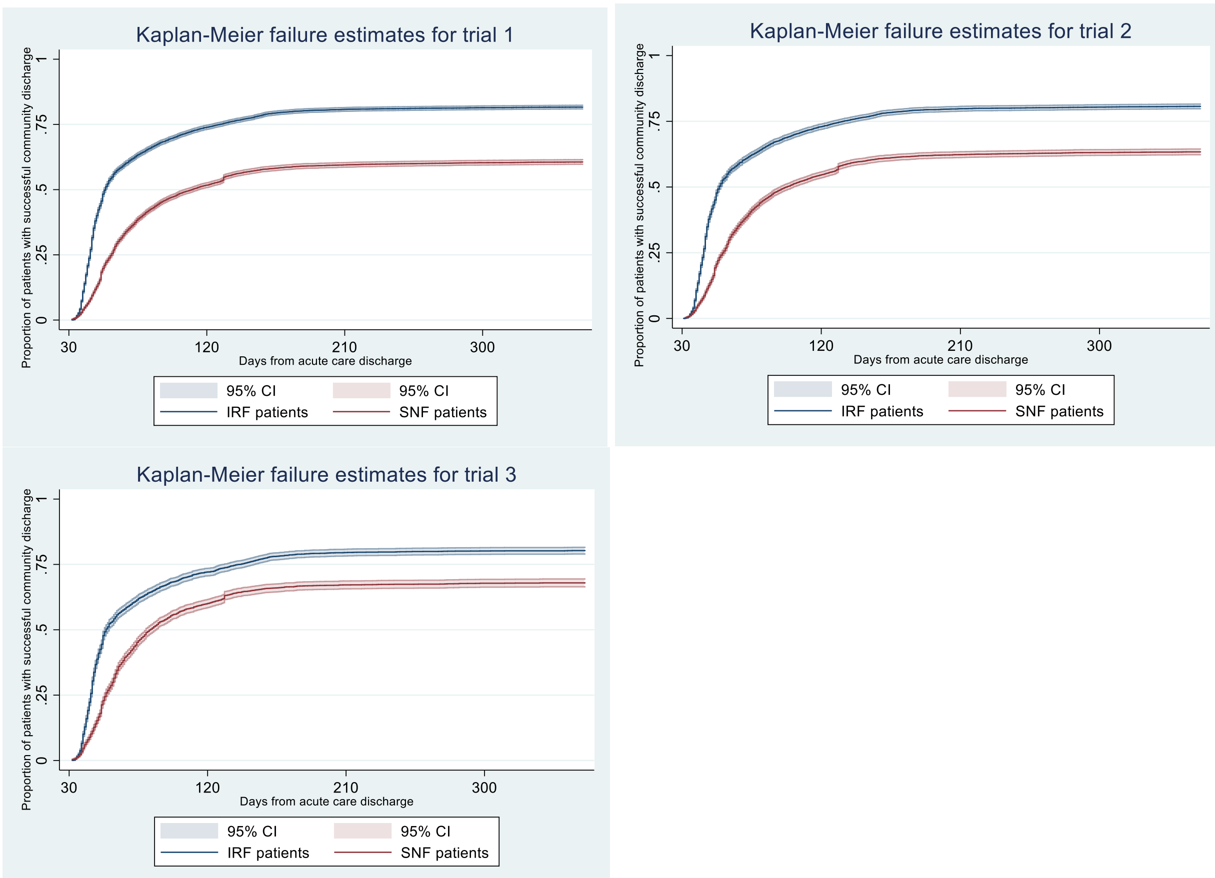
Table 4.4: Comparative binary outcomes for the three propensity score matched target trials comparing stroke rehabilitation at an Inpatient Rehabilitation Facilities compared to a Skilled Nursing Facilities

	Trial 1		Trial 2		Trial 3	
	Unadjusted (n=44,950 patients)	Matched (n=11,784 pairs)	Unadjusted (n=34,444 patients)	Matched (n=7,578 pairs)	Unadjusted (n=19,161 patients)	Matched (n=3,728 pairs)
1-year successful community discharge (95% CIs)						
Risk difference (95 % CI)	0.34 (0.33, 0.35)	0.21 (0.20, 0.22)	0.32 (0.31, 0.33)	0.17 (0.16, 0.19)	0.27 (0.20, 0.23)	0.12 (0.10, 0.14)
Relative risk (95 % CI)	1.82 (1.79, 1.85)	1.35 (1.32, 1.37)	2.02 (1.97, 2.07)	1.27 (1.25, 1.30)	1.98 (1.91, 2.05)	1.18 (1.15, 1.21)
Odds ratio (95 % CI)	4.41 (4.23, 4.63)	3.02 (2.83, 3.22)	3.83 (3.64, 4.03)	2.54 (2.34, 2.74)	3.15 (2.94, 3.38)	1.98 (1.77, 2.21)
E-value (lower bound of 95% CI)	N/A	2.04 (1.97)	N/A	1.86 (1.81)	N/A	1.64 (1.57)
1-year all-Cause Mortality (95% CIs)						
Risk difference	-0.25 (-0.26, -0.24)	-0.12 (-0.13, -0.11)	-0.25 (-0.26, -0.24)	-0.11 (-0.12, -0.09)	-0.22 (-0.24, -0.21)	-0.08 (-0.10, -0.06)
Relative risk	0.64 (0.63, 0.65)	0.85 (0.84, 0.86)	0.57 (0.56, 0.59)	0.86 (0.85, 0.87)	0.56 (0.54, 0.59)	0.89 (0.87, 0.92)
Odds ratio	0.34 (0.33, 0.36)	0.53 (0.51, 0.56)	0.36 (0.34, 0.37)	0.57 (0.52, 0.61)	0.39 (0.36, 0.42)	0.63 (0.57, 0.71)

Unadjusted: All available patients for each trial
Successful community discharge=Discharge home and remained alive and outside of acute care, an inpatient rehabilitation facility or a skilled nursing facility
Abbreviations CI: Confidence interval
For match pairs: standard errors used to calculate 95% CIs adjusted to account for dependence between pairs
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

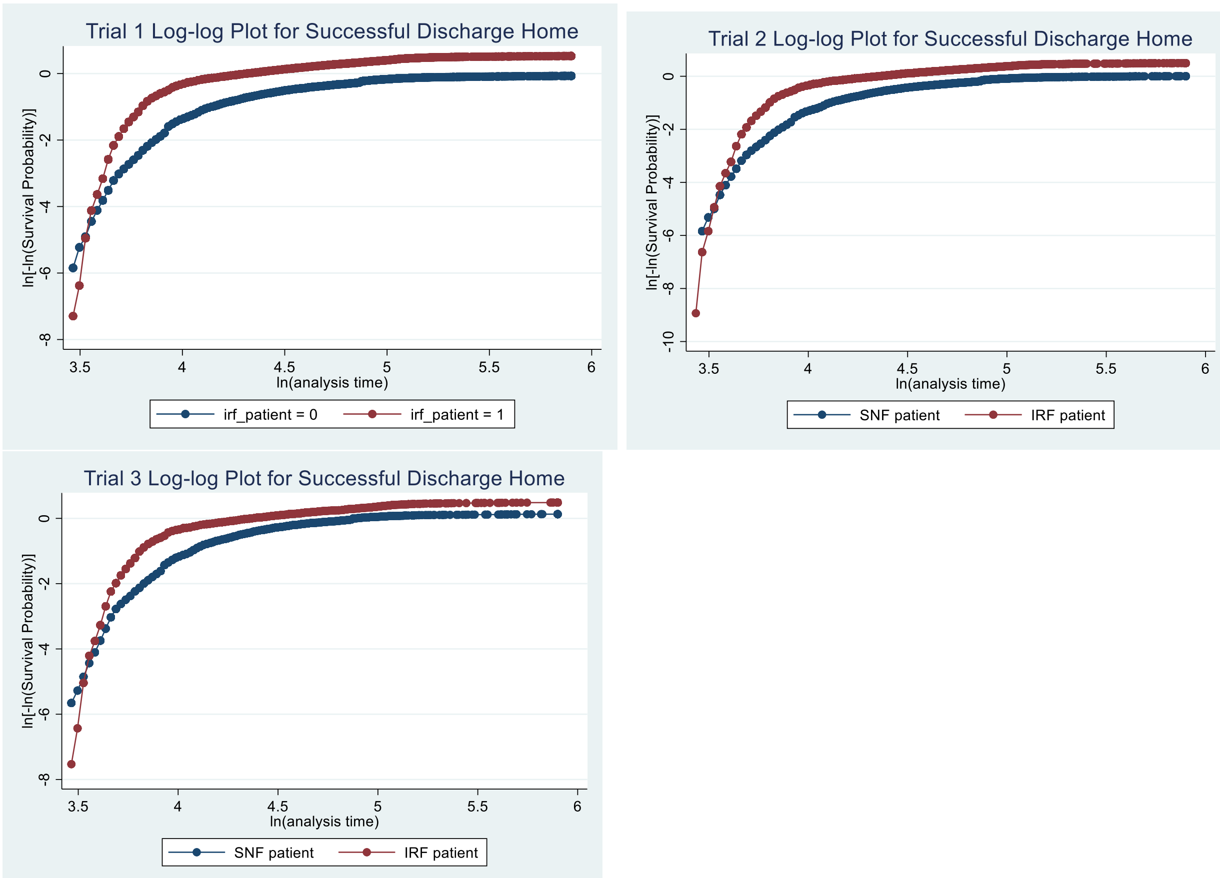
Comparative outcomes for time-to-event endpoints

Figure 4.5 shows the Kaplan Meier failure curves stratified by treatment arm (IRF vs SNF) for time-to-successful community discharge for each matched target trial. Initially, IRF patients achieve successful community discharge at a much faster rate compared to SNF patients. However, after about 150 days both curves flatten and become near parallel indicating that very few additional patients achieve successful community discharge after this timepoint. The average relative difference in these curves over the 1-year follow up was quantified by a HR of 1.99 (95% CI: 1.93, 2.05) (Table 4.5). The treatment effect for IRFs, remained significant but was smaller in trial 2 (HR: 1.80 (95% CI: 1.73, 1.87)) and 3 (HR: 1.57 (95% CI: 1.49, 1.65)) (Table 4.5). However, care should be taken when interpreting the HRs because based on the Schoenfeld global test, the proportionality assumption was not met ($p<0.05$). However, upon visual inspection of the log-log survival plots, there was only minor overlap very early during the follow up period, with the rest of the curves roughly parallel (Figure 4.6). For the observed (Kaplan Meier estimate) vs. predicted (Cox model estimate) survival curves (inverse of the failure curve), there were relatively minor differences which indicates that a reasonable degree of proportionality was present (Figure 4.7). In the sensitivity analysis, in which death was treated as a competing risk, the cause-specific HR was 1.84 (95% CI 1.81, 1.87) for trial 1, and 1.67 95% CI: (1.64, 1.71) and 1.50 (95% CI: 1.45, 1.55) for trials 2 and 3 respectively (Table 4.5). Thus, accounting for death resulted in HR estimates that were 10-15% lower.



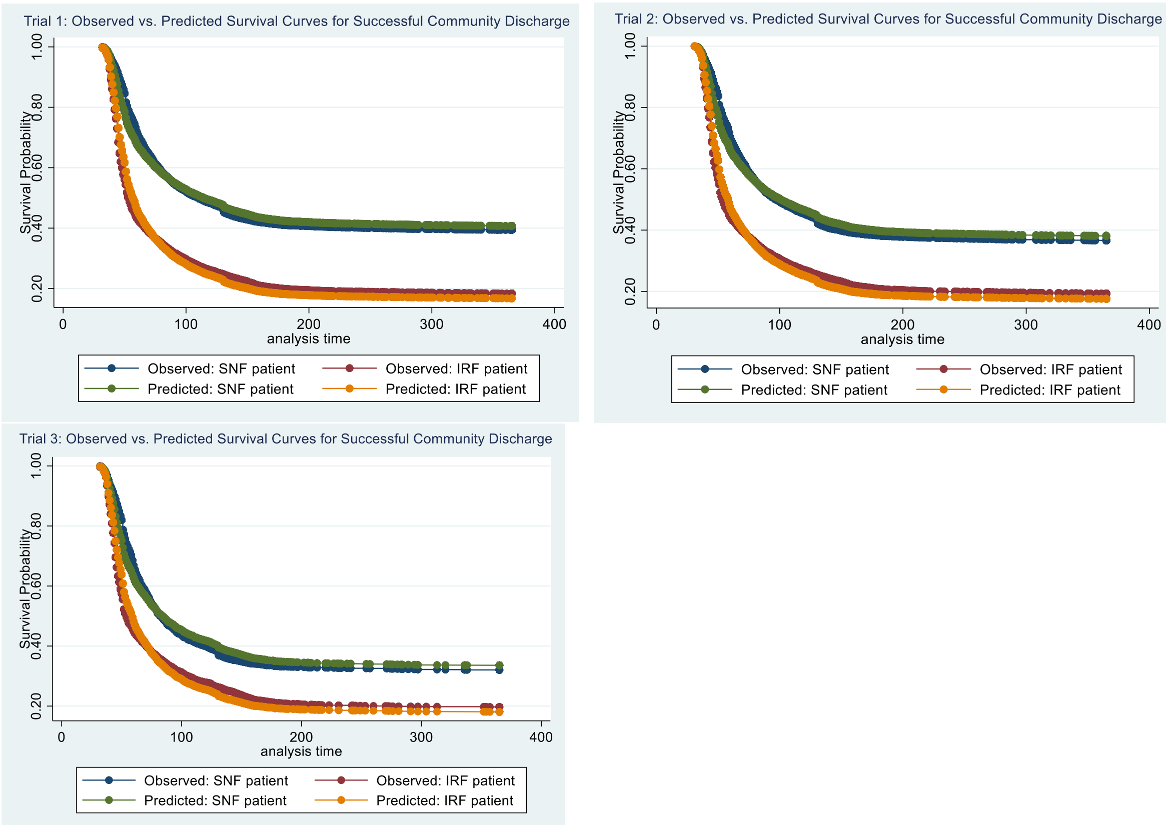
Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.5: Kaplan Meier failure curves for 1-year successful community discharge following rehabilitation at an Inpatient Rehabilitation Facility of a Skilled Nursing Facility among acute stroke patients



Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.6: Log-log plots of successful community discharge *failure* curves used to assess proportionality assumption for Cox proportional hazards model



Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.7: Observed (Kaplan-Meier estimate) vs. Predicted (Cox model estimate) *survival* plots for successful community discharge used to assess proportionality assumption for Cox proportional hazards model

Table 4.5: Hazard Ratios and 95% CIs for comparative time-to-event outcomes for the three propensity score matched target trials that compare stroke rehabilitation at Inpatient Rehabilitation Facilities (IRFs) compared to Skilled Nursing Facilities (SNFs)

Primary Analysis: Cox Proportional Hazard model						
	Trial 1 (95% CIs)		Trial 2 (95% CIs)		Trial 3 (95% CIs)	
	Unadjusted	Matched	Unadjusted	Matched	Unadjusted	Matched
Successful community discharge	2.57 (2.51, 2.62)	1.99 (1.93, 2.05)	2.30 (2.24, 2.37)	1.80 (1.73, 1.87)	2.02 (1.95, 2.10)	1.57 (1.49, 1.65)
All-cause mortality	0.39 (0.38, 0.40)	0.58 (0.55, 0.61)	0.40 (0.39, 0.42)	0.61 (0.58, 0.65)	0.43 (0.21, 0.46)	0.68 (0.62, 0.75)
Competing Risks Model (death=competing risk)						
Successful community discharge	2.25 (2.20, 2.30)	1.84 (1.81, 1.87)	2.02 (1.97, 2.07)	1.67 (1.64, 1.71)	1.78 (1.72, 1.84)	1.50 (1.45, 1.55)

* Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF

Match: Patients were matched based on their probability of discharge to an IRF (vs. SNF) which was estimated from a single level logistic regression model

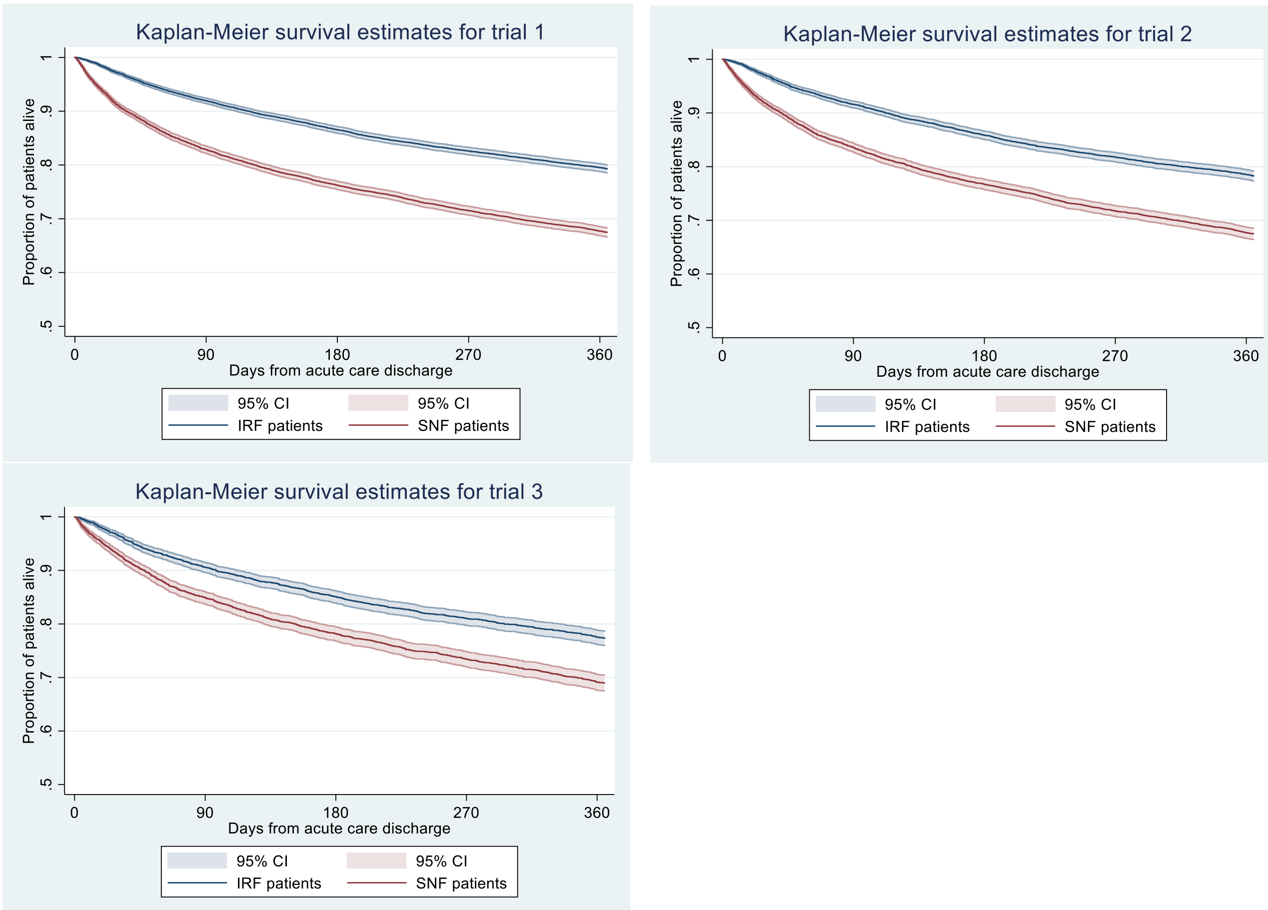
Unadjusted: All available patients for trials 1 (n=44,950), 2 (n=34,444), and 3 (n=19,161)

Trial 1: All matched patients

Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients

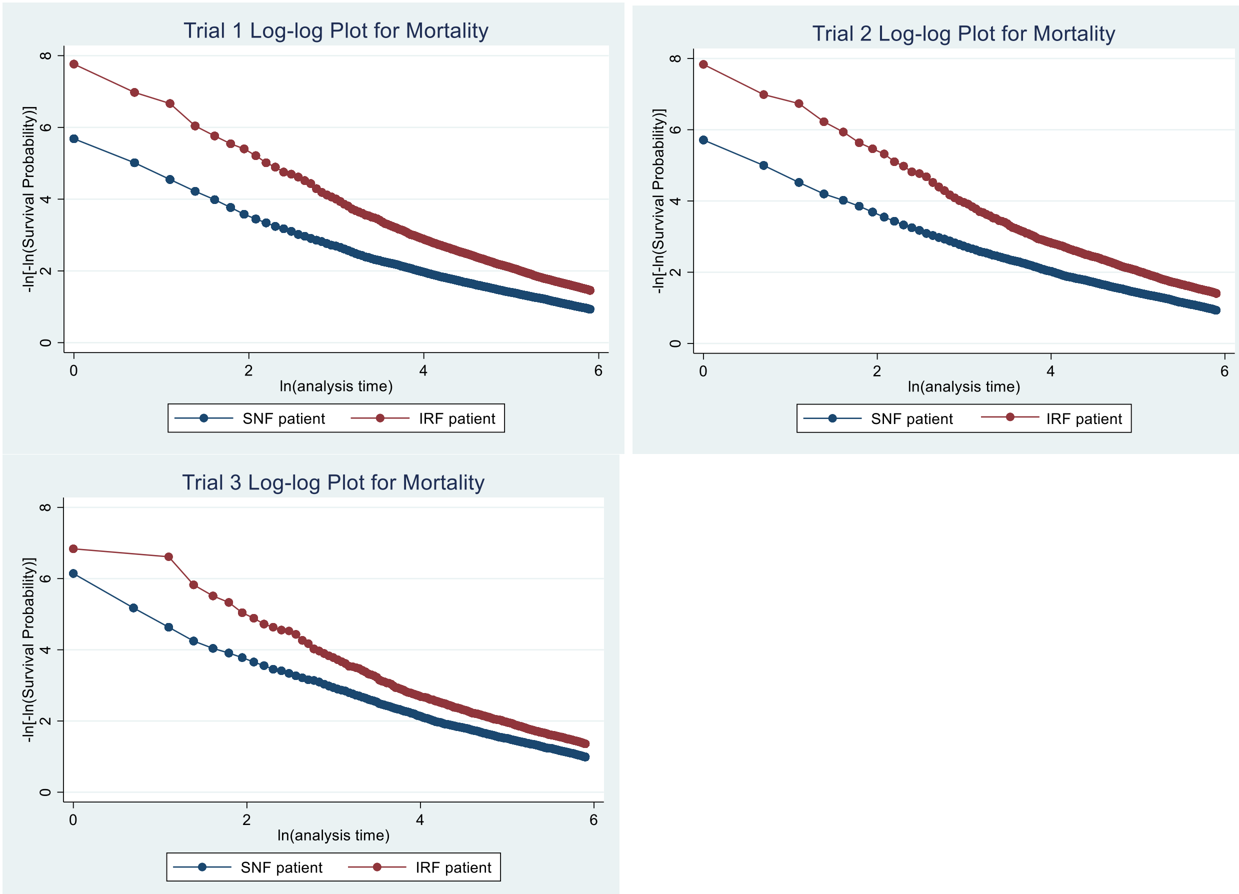
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.8 shows the corresponding Kaplan Meier survival curves for all-cause mortality. Unlike successful community discharge, there was no plateau as patients continued to die throughout the 1 year of follow up. The HR for the matched trials 1, 2, and 3 were 0.58 (95% CI:0.55, 0.61), 0.61 (95% CI: 0.58, 0.46) and 0.68 (95% CI: 0.62, 0.75) respectively, indicating that IRF care was associated with between a 32 (Trial 3) to 42% (Trial 1) lower risk of death (Table 4.5). Similar to successful community discharge, proportionality was not statistically met by the Schoenfeld global test. Visual inspection of the log-log plots (Figure 4.9) and observed vs. predicted survival curves (Figure 4.10) show that there was less proportionality during the first few months after acute care discharge, but that there was a reasonable degree of proportionality present overall.

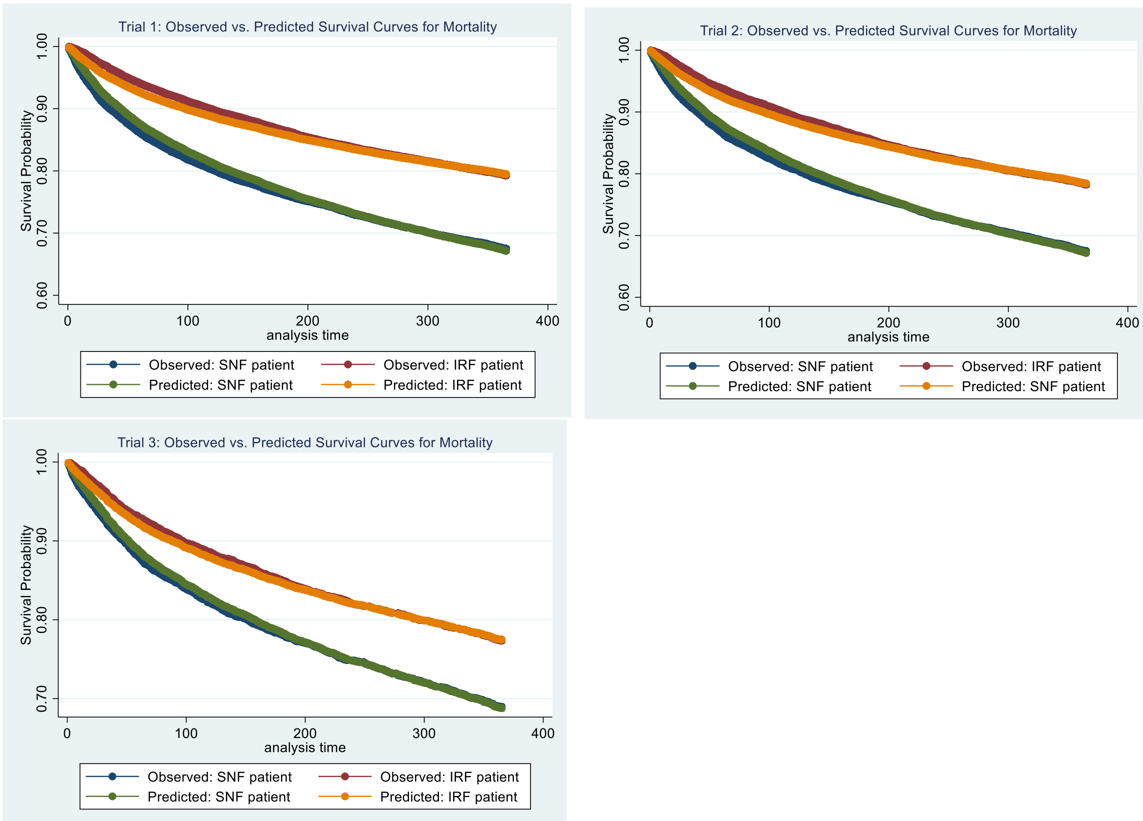


Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.8: Kaplan Meier survival curves for 1-year all-cause mortality following rehabilitation at Inpatient Rehabilitation Facilities vs. Skilled Nursing Facilities Among Acute Stroke Patients



Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients
Figure 4.9: Log-log plots for 1-year all-cause mortality used to assess proportionality assumption for cox proportional hazards model

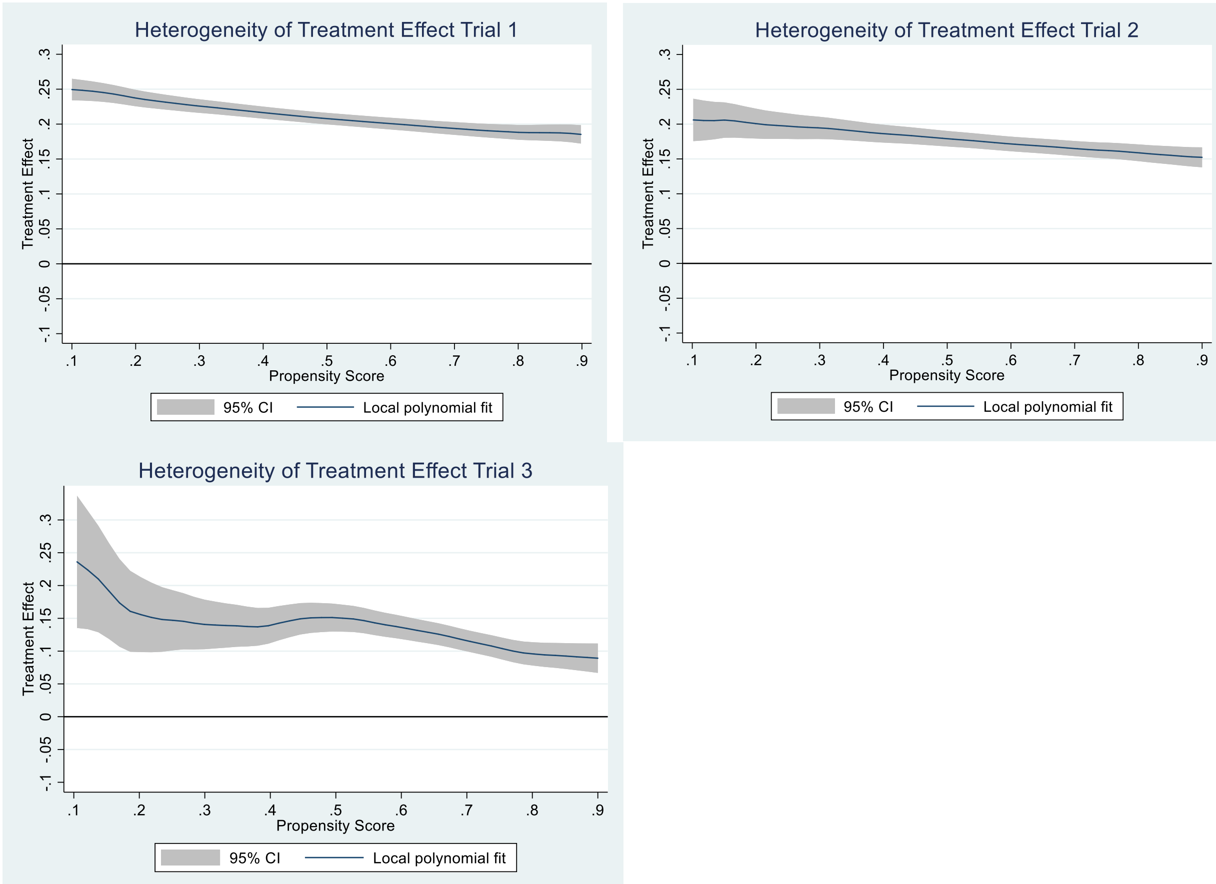


Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.10: Observed (Kaplan-Meier estimate) vs. Predicted (Cox model estimate) *survival* plots for 1-year all-cause mortality used to assess proportionality assumption for cox proportional hazards model

Heterogeneity of treatment effect

Figure 4.11 depicts the estimated treatment effect for IRF compared to SNF care by the estimated PS for the three trials. The local polynomial curve with its 95% CI is a non-parametric regression of the risk difference in successful community discharge (yes/no) for each matched pair (i.e., IRF patient outcome 1/0 -SNF patient outcome 1/0) over the range of the estimated PS. The polynomial curve shows the estimated treatment effect (estimated using the risk difference) in the proportion of patients who were successfully discharged home. For trial 1, the treatment effect for IRF patients was 0.25 for the patients who were least likely to go to an IRF (PS of 0.1). This indicates that if 100 SNF patients had received care at an IRF then an additional 25 of these same patients would have been successfully discharged home. However, among patients who had equally high likelihood to go to an IRF (PS of 0.9) the RD of treatment was only 0.18, indicating that for patients most likely to be discharged to receive care at an IRF (vs. SNF), the treatment was 7% less effective on an absolute scale. The heterogeneity of treatment effect around the PS was most profound in trial 3, where the treatment effect (as measure by the RD) was over 2.5 times larger for patients with a PS of 0.1 (RD=0.24) compared to patients with a PS of 0.9 (RD=0.09). Again, indicating that the treatment effect was greatest for the patients who were least likely to receive it. However, there was more variation around this estimate which is shown by a non-linear line and wider 95% CIs and the overall RD for trial 3 was lower that trials 1 and 2.



Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
Treatment effect=IRF patient outcome (1 or 0) – SNF patient outcome (1 or 0)
Trial 1: All matched patients, Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients, Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Figure 4.11: Risk difference (treatment effect) in successful community discharge between matched Skilled Nursing Facility patients and Inpatient Rehabilitation Facility patients over the estimated propensity score

Sample size estimates

Table 4.6 shows the range of estimated sample sizes that would be needed to conduct a superiority trial that compared differences in the probability of successful community discharge between IRF vs. SNF care for stroke rehabilitation using a one-sided test with α<0.05 and power set at either 80 or 90%.[136] The range of estimated total sample sizes for the three emulated trials with RD estimates ranging from 21% to 12% was 114 (trial 1) to 330 (trial 3) with power set at 80%. With power set at 90% these estimates were 156 (trial 1) to 454 (trial 3) We also estimated the sample size that would be needed to detect a much lower treatment effect i.e., 5% and 2.5% difference using the same baseline success rate of SNF patients (observed in trial 3 (68%). A sample size of 2,056 (power set at 80%) and 2,846 (power set at 90%) patients would be needed to detect a 5% difference, while 8,424 (power set at 80%) and 11,668 (power set at 90%) patients would be needed for a 2.5% difference in successful community discharge.

Table 4.6: Sample size calculations for a superiority trial that compares the difference in 1-year successful community discharge which compares stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

Proportion of SNF patient success	Risk Difference	Total sample size (80% power)	Total sample size (90% power)
0.61 (Trial 1)	0.21	114	156
0.63 (Trial 2)	0.17	172	236
0.68 (Trial 3)	0.12	330	454
0.68	0.05	2,056	2,846
0.68	0.025	8,424	11,668

Abbreviations: IRF: Inpatient Rehabilitation Facilities, SNF: Skilled Nursing Facilities
*Sample size estimates for all trials were 1-sided with α set at 0.05
Trial 1: All matched patients
Trial 2: Matched patients treated a rehabilitation facility that received greater than 5 stroke patients
Trial 3: Matched patients treated a rehabilitation facility that received greater than 10 stroke patients

Sensitivity analysis results

We performed a sensitivity analysis by matching across hospitals (rather than within) using the starting population from trial 1 (i.e., 44,950 patients from 441 hospitals and treated at

745 IRFs and 5,974 SNFs). In the original analysis (Trial 1) when matching within facilities, 11,784 pairs were identified who were treated at 662 IRFs and 4,579 SNFs. When matching across facilities, 14,703 pairs were matched and these patients were treated at 701 different IRFs and 5,005 SNFs. As with the original analysis, all covariates were well balanced between the two groups as the ASDs were all <0.1 (See Supplemental Figure 4.1). Table 4.11 shows the descriptive outcomes for IRF and SNF patients both for trial 1 and the sensitivity trial population. Overall, matching across (rather than within) facilities had almost no effect on the occurrence of any outcome as all differences changed less than 1%.

Table 4.7: Descriptive outcomes for the sensitivity matched target trials comparing stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

	Trial 1 (n=11,784 pairs)		Sensitivity trial (n=14,703 pairs)	
	IRF (%)	SNF (%)	IRF (%)	SNF (%)
Successful community discharge				
90 day	68.1	45.2	68.5	45.3
1 year	81.6	60.6	81.9	60.9
All-cause mortality				
30 day	3.0	9.2	2.9	8.9
90 day	8.1	17.2	7.9	16.7
1 year	20.7	32.5	20.5	32.5
All cause acute hospital readmissions				
30 day	13.4	16.1	13.3	16.0
90 day	25.9	28.9	26.0	28.8
1 year	48.6	49.2	48.5	49.2

Trial 1: Patients matched within hospitals, Sensitivity Trial: Patients matched across hospitals
Time starts upon discharge from the acute care hospital
Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
For match pairs: standard errors used to calculate 95% CIs adjusted to account for dependence between pairs

Table 4.8 shows the outcome comparisons between IRF and SNF patients following PS matching for trial 1 and the sensitivity trial population. Similar to the descriptive outcomes, matching across (rather than within) hospitals had virtually no effect on the estimated treatment effect of IRF (vs. SNF) rehabilitation.

Table 4.8: Comparative binary outcomes for propensity score matched target trial #1 and the sensitivity trial both of which compares stroke rehabilitation at Inpatient Rehabilitation Facilities compared to Skilled Nursing Facilities

	Trial 1 (n=11,784 matched pairs)	Sensitivity trial (n=14,703 matched pairs)
1-year successful community discharge (95% CIs)		
Risk difference	0.21 (0.20, 0.22)	0.21 (0.20, 0.22)
Relative risk	1.35 (1.32, 1.37)	1.34 (1.32, 1.36)
Odds ratio	3.02 (2.83, 3.22)	2.94 (2.78, 3.12)
1-year all-Cause Mortality (95% CIs)		
Risk difference	-0.12 (-0.13, -0.11)	-0.12 (-0.13, -0.11)
Relative risk	0.85 (0.84, 0.86)	0.85 (0.84, 0.86)
Odds ratio	0.53 (0.51, 0.56)	0.52 (0.49, 0.55)

Trial 1: Patients matched within hospitals, Sensitivity Trial: Patients matched across hospitals
Time starts upon discharge from the acute care hospital
Successful community discharge=Discharge home and remained alive and outside of acute care, an IRF or a SNF
For match pairs: standard errors used to calculate 95% CIs adjusted to account for dependence between pairs

DISCUSSION

We used administrative data to emulate three target trials to estimate the treatment effect of rehabilitation at IRFs compared to SNFs in a population of fee-for-service Medicare beneficiaries hospitalized with acute stroke. Overall, our results are largely consistent with previous observational studies that show that acute stroke patients who were treated at IRFs have superior outcomes relative to discharge home and mortality compared to patients treated at SNFs.[10,37,120–122] Specifically, in our PS matched trials we showed that on a relative scale IRF patients were 18-35% more likely to be successfully discharged home, and 11-15% less likely to die within 1 year of discharge from the acute care hospital compared to SNF patients. Compared to prior comparative effectiveness studies, our analysis was conducted among a carefully selected subset of hospitals, patients, and rehabilitation facilities that we believe represent the ideal target population for a subsequent RCT. Additionally, we showed that the treatment effect size was attenuated (but not eliminated) when larger, more frequently used SNFs were compared to IRFs – as reflected by the results of trial 3. We also showed that the treatment effect of IRF care was the highest for the patients least likely to receive IRF care, which might indicate that negative selection may be present.[137] Finally, we used the results to conduct several sample size calculations to estimate that a subsequent superiority trial would need to have between 114 to 330 total patients (with power set at 80%).

Comparisons with previous studies:

Overall, our results are generally similar to previous comparative effectiveness studies that used Medicare data showed that acute stroke patients who were treated at IRFs have better outcomes compared to patients treated at SNFs[10,37,120–122] Our overall 1 year mortality rates were virtually identical to the 1 year all-cause mortality for IRF (17.9% vs 17.9%) and SNF (38.8% vs. 38.6%) rates reported from a study of 69,212 Medicare stroke patients who were treated at 1,146 hospitals that participated in the stroke Get With The Guidelines cohort study.[97] This study used inverse propensity weights and estimated the HR for 1 year all-cause mortality to be 0.65, which was consistent with the HRs that were estimated for trials 1 (0.58) and 3 (0.68). We were the first study to model 1-year successful community discharge comparing IRF vs. SNF care for stroke patients, although a 2006 study used a multivariable logistic regression model to discharge home for IRF and SNF patients.[10] This study by Deutsch, et al linked clinical data for 58,724 Medicare beneficiaries with stroke and reported that generally, patients treated at IRFs had around a 2-fold increase in the odds of being discharged home (but the exact effects depended on the specific disability strata) .[10] This estimate corresponds most closely to our OR estimate of 1.98 from trial 3. Similar to trial 3, the Deutsch, et al study included a selective subset of larger, high performing SNFs that measured patient activity level function with the FIM™ Instrument rather than the Minimum Data Set.[10]

Novel findings:

Overall this study provided at least two novel findings. First, by conducting three emulated trials we identified that the relative treatment effect of IRF vs. SNF care is highly dependent on the types of SNFs that were compared. We could only observe that patients at large, frequently used SNFs had better outcomes (compared to smaller SNFs). Unfortunately, data on facility-level processes of care (i.e., the type, frequency and intensity of care) are poorly characterized and understood for rehabilitation facilities.[7,8,112] Recently, several studies have sought to characterize facility-level variation for IRFs, but equivalent studies for SNFs are currently lacking.[23,107,108] In the United States, there are over 12 times the number of SNFs compared to IRFs and the diffuse nature of SNF care presents a major methodological challenge to accurately profile the very large number of relatively small facilities.[21] Our study was not designed to characterize variation or drivers in care quality in SNFs. However, the SNFs in trial 3 (which represent 6.9% of all SNFs) are not representative of typical SNFs because they were much larger, more frequently used and likely represent the best of SNF care.

Second, if true, our finding that heterogeneity of treatment effects was present across the PS is provocative. This finding indicates that the patients who were least likely to be treated at an IRF, experienced the largest treatment benefit when they received this care which represents a type of negative selection bias.[137] National data shows that patients with characteristics associated with favorable recovery following stroke (e.g., younger age, fewer comorbidities) were more likely to be discharged to receive care at IRFs.[15–17,41] Under the assumption that these patients have favorable recovery trajectories, the high intensity care provided by these facilities may have smaller effects on shifting these trajectories (compared to patients who are sicker and have lower activity level function).[8,11] However, our results are speculative and more work is needed to validate this finding and the result should be interpreted with two considerations in mind. First, the outcome was binary was a proxy variable for function which provides a low ceiling effect for activity level function as there is a large range of function beyond the minimal threshold that patients need to live at home. Second, the observed heterogeneity could be caused by unmeasured variables.[133,137]

Potential explanations for why IRF patients do better:

Several clinical reasons may explain why patients discharged to an IRF had improved outcomes relative to patients discharged to SNFs. First, there could be a selection bias that we were unable to adjust for in that healthier patients with favorable recovery trajectories are more likely to be discharged to an IRF. Second, patients treated at IRFs receive a much larger therapy dose in the early stage of their recovery period and this therapy is provided by a highly specialized, multidisciplinary rehabilitation treatment team. Early initiation of intense rehabilitation therapy and care is provided by multidisciplinary treatment teams are both associated with improved physiological and activity level functional outcomes.[33,34] Third, patients at IRFs have closer clinical monitoring and have greater access to physicians and nurses.[4] In addition, IRFs are often physically integrated into other hospitals which could provide greater access to other medical specialists as well as diagnostic and treatment technologies.[37] Finally, around 15% of SNF patients transitioned to become long term care nursing home patients. It is conceivable that exposure to the nursing home and comfort with the nursing home staff could ease the transition to long term nursing home care compared to moving in from another facility (be it SNF or IRF) or home.

Implications for the design of a subsequent pragmatic RCT:

Our results provide important information to inform the design of the subsequent trial. First, by showing the range of effect estimates across the trials we show the importance for the need to carefully consider which IRFs and especially SNFs should be compared. In the United States, the Patient Centered Outcomes Research Institute (PCORI) is a major funder and advocate for pragmatic RCTs because of their ability to rapidly fill real world evidence gaps and because their results can be quickly translated to affect the current clinical care landscape.[138] According to their methodological standards, trial 1 would be discouraged because comparisons of "usual care" are often ill-defined and highly variable.[139] Second, we calculated a range of sample sizes based on the estimates of effect size for IRF vs. SNF care from the three trials for the probability of 1-year successful community discharge. However, successful community discharge is a crude proxy measure of function and the actual trial would use a primary outcome measure that were able to capture patient function across the bio-psycho-social-environmental domains identified by the ICF model.[29] Examples of potential measures could include the Activity Measure of Post-Acute Care (AM-PAC), 16 item Stroke Impact Scale, or the modified Rankin score.[94,140] For this study, we used successful community discharge because other common outcomes (e.g., rehospitalization and death) do not measure function well, and the functional measures used in IRFs (Inpatient Rehabilitation Facility-Patient Assessment Instrument) and SNFs (Minimum Data Set) are not directly comparable because separate assessment measures are used and the data is collected at different time points.[8,141,142,28] Third, our follow-up time was 1-year but a future study trial may consider follow-up times of 6 months because the Kaplan –Meier survival curves flattened out by around 150 days. Finally, we calculated sample sizes for a superiority trial and identified that for the emulated trials a sample size (power set at 80%) of between 114-330 patients would be needed for trials 1 and 3 respectively.

Study strengths

A major strength of this study is that we emulated our target trials within a pre-identified target trial population that controlled for the influence that hospital contextual effects have on patients which may bias effect size estimates that are based on comparing average outcome differences between IRF and SNF populations. Overall, within the emulated trial framework, we captured a treatment effect size estimate directly from the population of interest and our results can are interpreted within the same counterfactual framework as an RCT (i.e., ATT estimate).[51] In addition, by emulating three trials we were able to provide a range of effect size estimates to inform trial design decisions. In addition, our data is nationally representative for Medicare patients, so our results have excellent generalizability for older Americans. Finally, in our sensitivity analysis, we showed that there are unlikely to be unmeasured clinical selection forces which operate across hospitals as there was virtually no difference in the comparative effectiveness estimate observed by matching across (rather than within) hospitals (i.e., Trial 1 vs. the sensitivity trial).

Study limitations

However, our results should be interpreted with several important methodological limitations in mind. First, the use of successful community discharge as our primary outcome may have biased the results towards favoring IRF patients due to a violation of the ignore-ability assumption of a PS analysis (i.e., that treatment assignment is independent of the outcome).[51] This assumption may have been violated because potential for community discharge is a clinical indication for IRF referral.[4] Ultimately, successful community discharge is a composite measure where "success" is contingent on some combination of patient activity level function, social support (e.g., availability of informal care) and environmental demands (e.g., physical challenges of the home environment).[8,11] Unfortunately, we were unable to control for many of these factors so prominent residual confounding likely persists. However, despite these limitations successful community discharge was developed by CMS as a proxy for a common patient centric measure of function to be used across all types of rehabilitation settings.[127] Second, as with all observational data analyses adjustment for systematic differences between populations assumes valid measurement of confounders. However, our use of administrative claims data relied on several measures of health service use as proxies for medical acuity. Although our adjustments reduced mortality differences between the two populations, the significantly higher mortality rate for SNF patients (especially at 30 days) likely indicates the presence of strong residual confounding from unmeasured factors.[28] By calculating E-values we showed that an unmeasured confounder with a RR of between 2.04 (trial 1) and 1.64 (trial 3) would be needed to nullify the results. It is entirely plausible that one or more unmeasured factors such as post-stroke function or social support could have a RR of this size.[41] Third, the use of a narrow caliper (set at 0.01), matching within facilities, and matching without replacement limited the number of matched patients in our sample. However, we still had large numbers of matched pairs who were treated at a large number of hospitals and rehabilitation facilities and these decisions improve internal validity.[123,143]

In conclusion, our observational analysis showed that among a carefully selected target trial population, acute stroke patients discharged to IRFs were more likely to be successfully discharged home and less likely to die within 1 year. We also showed that the magnitude of this treatment effect was conditional on the types of facilities being compared – particularly SNFs, and to patients' baseline propensity for receiving IRF care. While our populations were well balanced with respect to measured confounders, several potentially important unmeasured confounders (e.g., social support, baseline function, home environment) were not available. Despite these limitations, we showed that stroke patients who received rehabilitation at large frequently used SNFs (Trial 3) have outcomes that were closer to those of IRF patients but still meaningfully different. An RCT would clarify these differences because random patient allocation would facilitate equal distributions of both measured and unmeasured confounders resulting in a more valid comparison.

CHAPTER 5: GENERAL DISCUSSION

OVERVIEW

Every year around 800,000 people suffer a stroke an after a hospital stay of a few days, around half of stroke patients will be discharged to either IRFs or SNFs.[5,6,96] However, whether stroke outcomes are better following rehabilitation at IRFs or SNFs is unknown because data is limited to observational designs and the selection forces for IRF and SNF care are both complex and strong.[15,144] Thus, a RCT is needed to answer this question, however the design of such a study is complicated by several practical and ethical issues. The purpose of this dissertation was to use Medicare claims data to inform the design and emulate such a clinical trial. We achieved this through a series of three aims. First, we aimed to identify drivers of IRF or SNF care. We identified several patient- (e.g., age, sex, dementia) and hospital-level (e.g., had an affiliated IRF unit, urban setting) factors which were associated with discharge to an IRF (vs. SNF) and that overall hospitals were responsible for around 30% of the variation in IRF and SNF use. However, we also showed that there was substantial heterogeneity of hospital effects on influencing IRF or SNF discharge and that half of patients attended an IRF or a SNF favoring hospital that changed their predicted probability of IRF discharged by over 10%. Second, we assessed the effect that several hospital level inclusion criteria had on patient level generalizability in order to identify the target trial population that we believed afforded an optimal pragmatic-explanatory balance for the subsequent emulated RCT.[48,56] To identify this population, we profiled hospitals based on their propensity to discharge stroke patients to IRFs (vs. SNFs) and inferred referral networks by examining the number and type of rehabilitation facilities that patients were discharged to. Our target trial population included 44,950 patients (30.8% of the starting sample) who were treated at 441 hospitals (14.5% of hospitals) and were subsequently discharged to 745 IRFs (64.8% of IRFs) and 5,974 SNFs (48.2% of SNFs). Third, we used a matched propensity score analysis to emulate three clinical trials that differed in the frequency of utilization of the IRFs and SNFs that were included. Overall, on a relative basis we showed that patients who were treated at an IRF were between 18-35% more likely to be successfully discharged home (i.e., discharged home and remained alive at home for >30 days) and were between 11-15% less likely to die within one year of discharge from the hospital. The different trial effect size estimates in Trials 2 and 3 were primarily driven by improved outcomes for SNF patients who were treated in larger, more frequently used SNFs compared to SNF patients treated at smaller, less frequently used SNFs which were more common in the Trial 1 population. Overall, our results were limited by the inability to adjust for important unmeasured confounders (e.g., social support, home environment) so it is unclear how much of the observed difference was due to residual confounding.[28] However, by calculating E values we showed that a moderately sized unmeasured confounder with a RR of 2.04 (Trial 1) and 1.64 (Trial 3) would be enough to nullify the observed differences between IRFs and SNF patients. An RCT would eliminate such biases and provide a more valid comparison.

SUMMARY OF THE OVERALL FINDINGS

In the United States, stroke is the 5th leading cause of death and the leading cause of adult disability.[1] A typical hospitalization for an acute stroke patient only lasts a few days and the primary focus is on medical stabilization.[24,97] However, most patients continue to require PAC to address residual disabilities and medical needs. Around half of stroke patients will receive PAC at either an IRF or SNF.[5,6] IRFs provide time-intensive therapy under regulations that specify minimum clinical and administrative requirements, whereas SNFs provide moderately intensive therapy delivered by nurses and other rehabilitation professionals but without direct physician supervision.[4] Clinically, IRF care is indicated for high acuity patients who are expected to have significant physiological and activity level functional recovery gains and be discharged back to the community. Whereas SNF care is indicated for a broader range of patients who are expected to make only partial recovery.[4] However, despite these stated differences in clinical indications there is substantial variation in the type of patients who are discharged to receive IRF and SNF care as well as processes of care (i.e., the type, frequency, and intensity of care) both between and within these two settings.[15–17] This variation has significant financial implications as variations in PAC use is the largest driver of regional variation in Medicare spending.[26]

Previous comparative effectiveness studies for IRF vs. SNF care for stroke patients generally found IRF patients to have superior outcomes across a range of domains (i.e., activity level functional gain, community discharge, and mortality) compared to SNF patients (See Table 1.1 from Chapter 1).[10,37,97,120,121] However, the results of these observational studies are limited, as the results of several systematic-reviews have found that discharge to IRF and/or SNF facilities are influenced by a complex mix of patient (e.g., age, sex, comorbidities), hospital (e.g., has an affiliated IRF unit, for-profit status) and environmental factors (e.g., State).[144–146] In aim one, we sought to 1) identify patient and hospital level factors that were associated with discharge to an IRF (vs. SNF). 2) Evaluate general hospital contextual effects- (i.e., the degree to which the hospital influences patient level outcomes).[62,78] 3) Characterize the heterogeneity of hospital effects on individual predicted probabilities of IRF (vs. SNF) discharge. Consistent with previous studies, we identified that several sociodemographic (e.g., age, race, and sex), clinical (e.g., dementia) and health service utilization (e.g., tPA or gastrostomy tube use) all had moderate associations with IRF discharge.[15–17,41] In addition, we identified that several hospital-level characteristics (e.g., for-profit status, having an affiliated IRF unit) had large average associations with IRF discharge, but there was substantial variation in the magnitude and directions of these associations. We quantified the magnitude of general hospital contextual effects using ICCs and showed that in the unadjusted model, hospitals contributed 27% of the variation in IRF and SNF discharge (ICC=0.27). Interestingly, patient case-mix adjustment increased general hospital contextual effects (ICC=0.33). Finally, we risk stratified hospitals based on their propensity to discharge patients to an IRF or SNF. By stratifying, we identified that for around half of Medicare acute stroke patients who attended IRF and SNF favoring hospitals, hospitals directly changed the predicted probability of IRF discharge by over 10% for over 80% of their patients.

The design of any trial should be guided by the relevant causal question of most interest to patients, clinicians, and policymakers. The best design to provide unbiased real-world comparative estimates to inform such policies is a pragmatic clinical trial.[55,138] However, a key challenge when designing such a trial is how to optimize the relative pragmatic-explanatory balance.[56] This balance depends on the specific causal question of interest, but relates to the ability of the trial to address efficacy (i.e., explanatory) issues to explain whether and how the trial can work, with effectiveness (i.e., pragmatic) issues of whether the results apply to a broad range of patients.[48] In aim 2, we sought to identify a target trial population that optimized this balance by exploring the effect that hospital level inclusion criteria had on patient level generalizability. We first used a multi-level logistic regression model to identify hospitals with typical discharge patterns (based on not having statistically significant hospital random intercepts). We included these hospitals because we believed that these hospitals would be more likely to participate in a trial where 1:1 random patient allocation would not result in substantial changes from their current practices. Second, we explored the effect of hospital case volume by including hospitals with either >20, >50, or >100 acute Medicare stroke rehabilitation patients over the two-year study period. Finally, because in a subsequent RCT it would not be feasible to enroll a very large number of IRFs and SNFs, we explored the effect on only including hospitals that were part of either regular use referral triads (i.e., discharged at least 5 patients to both a single IRF *and* SNF over a two-year period) or frequent used referral triad (i.e., discharged at least 10 patients to both a single IRF *and* SNF over a two-year period). Using the three sequenced eligibility criteria, we identified a final target trial population that included 44,950 patients (30.8% of patients) who were treated at 441 (14.5% of hospitals) and subsequently discharged to 745 IRFs (64.8% of IRFs) and 5,974 SNFs (48.2% of SNFs). This target trial population was highly representative of the national Medicare acute stroke population, but target trial hospitals were very different (e.g., they were larger, more likely to be in an urban setting, more likely to be affiliated with a medical school). We subsequently used this population to emulate the desired RCTs in the subsequent study of the dissertation.

In aim 3, we used the target trial population identified in study two to emulate three pragmatic RCTs that compared stroke rehabilitation at IRFs compared to SNFs. Emulated trials are hypothetical RCTs in which observational data analysis mimics the design features of a true trial (e.g., explicit time zero (t_0) and synchronized treatment assignment).[49,50] We emulated randomization using a matched propensity score. The three trials included common patient- and hospital-level eligibility criteria, but different rehabilitation facility level (i.e., IRFs and SNFs) criteria: Trial 1: included all rehabilitation facilities from the starting trial population, Trial 2: included rehabilitation facilities that treated ≥5 patients, and Trial 3: included rehabilitation facilities that treated ≥10 patients. We replicated three trials because there are known differences in the quality, type, and intensity of rehabilitation care provided across rehabilitation facilities.[7,27]

Overall across the three emulated trials, we found that stroke patients treated at IRFs had superior outcomes compared to patients treated at SNFs. This result was largely consistent with previous studies.[10,37,97,120,121] More specifically, on a relative basis we showed that patients treated at IRFs were between 18% (Trial 3) to 35% (Trial 1) more likely to be successfully discharged back to the community by 1 year, and were between 11% (Trial 3) and 15% (Trial 1) less likely to die within 1 year of discharge from the acute hospital. The difference in the estimated effect sizes across the three trials was almost entirely driven by patients treated at large regularly used SNFs (Trial 3) having better outcomes compared to patients treated at small infrequently used SNFs (Trial 1). In addition, by calculating E-values we showed that a moderately sized unmeasured confounder (e.g., post-stroke function or social support) with a RR of 2.04 (Trial 1), 1.86 (Trial 2), or 1.64 (Trial 3) would be needed to nullify the observed differences between the IRF and SNF patient outcomes. In our sensitivity analysis, we conducted a separate trial in which we matched across (rather than within) hospitals. This differential matching method had virtually no effect on any of the outcomes, indicating that unmeasured clinical selection forces that are present across hospitals are unlikely to affect outcomes. Finally, we also showed that there was heterogeneity of treatment effect across the propensity score.[133] Specifically, we showed that patients who were least likely to be discharged to an IRF, had the largest relative benefit of IRF (vs. SNF) care. While the overall comparative effectiveness estimate was closest for Trial 3, the SNFs included in this trial only represented around 20% of all of the SNFs that treated stroke patients. Thus, careful consideration is needed towards which types of IRFs and SNFs should be compared to address the question that is of most interest to patients, clinicians, and policy makers.

SUMMARY OF RECOMMENDATIONS FOR A FUTURE TRIAL

Overall, the results from the series of three studies of this dissertation provides a range of information to inform the design of a subsequent pragmatic RCT. Table 5.1 shows the emulated trial protocol and potential modifications to the emulated protocol that could be considered for an actual trial. First, compared to the emulated trial, patients in the actual trial must first be identified as being in clinical equipoise. As discussed in aim two, clinical equipoise is the point for which there is genuine uncertainty towards which treatment would provide optimal care, and is essential to justify random treatment allocation.[113,147] Among previously published RCTs, a range of options have been used to identify patients in equipoise including, variations in clinical practice, lack of RCT data, and investigators declaration.[114] Many studies pre-select equipoise patients using tight inclusion/exclusion criteria, but two newer and novel approaches to identify patients in real time have been developed.[148,149] In one approach, a mathematical prediction model is developed and patients with equivalent predicted outcomes are flagged as being in mathematical equipoise.[149] Notably, this approach developed prediction models using data from previous RCTs, but unfortunately no such data is available for IRF vs. SNF care for stroke patients. In another approach, a small panel of experts independently reviewed the medical records for patients who met inclusion criteria for surgical vs. medical treatment. These experts then indicated how much they believed surgery would benefit the patient on a 7-point Likert scale. These scales were then pooled, and discordance was then statistically modeled. Patients with the most uncertainty were identified as being in equipoise.[148] Second, in the trial, randomization stratified by hospital should be considered, as this method would ensure balance between treatment groups for each hospital.[150] Third, as discussed in study three, careful consideration is needed to determine appropriate follow-up time. For stroke rehabilitation, this consideration should be driven by several considerations including the primary outcome selected and the specific question that the trial is designed to answer. Many stroke trials choose 90 days after stroke onset as the time that their primary-end point is assessed, and many patients' activity level functional recovery trajectories plateau around this time.[40] However, longer term follow-up times may better account for substantial heterogeneity among recovery trajectories and provide data for the maintenance of recovery gains.[8,11] The assessment of maintenance is crucial as the goal of rehabilitation is long term sustainable functional gain and it is important to control for length of stay differences between rehabilitation care at IRF and SNFs.[8] Fourth, as discussed in study three, selecting the primary outcome to address the specific question of interest is essential. In the emulated trial, we used successful community discharge as the primary outcome because other common outcomes (e.g., rehospitalization and death) do not measure activity level function, and the functional measures used in IRFs (IRF-Patient Assessment Instrument) and SNFs (Minimum Data Set) are not directly comparable.[8,141,142,28] However, CMS is implanting new Quality Reporting Program item sets for IRFs, SNFs, and Home Health to address this problem. For policymakers, successful community discharge is valuable because it captures a point at which patients stop using expensive health services (i.e., hospitals, IRFs, or SNFs).[127] However, there is a range in the level of physiological, activity, participation, and environmental level of function among people that live at home. For patients, measures that can better contextualize a broader range of function across all bio-psycho-social-environmental domains based on the ICF model) would be of greater value.[29] For example, one candidate measure would be the AM-PAC which is an easy to use patient-self assessment and is able to capture patient level activity function.[151,152] Other potential outcomes measures include Patient-Reported Outcomes Measurement Information System (PROMIS) based patient reported outcomes related to function, or quantity of life based assessments such as the 16-item Stroke-scale-impact-16. In the trial, it would be good to continue to capture successful community discharge (at 6 months) for a follow up study that could compare the results of the emulated trial with that of the actual trial, as well as other traditional legacy measures used in many stroke trials such as modified Rankin score. Finally, as discussed in aim 3, careful consideration is needed to decide between if the trial should be a superiority trial or if it should be a non-inferior trial. An superior trial would answer the question if outcomes for IRF were better than SNF care, while a non-inferior trial would answer the question that SNF care is at least not worse than IRF care.[136]

Table 5.1: Comparison of the design features of the emulated trials compared to *potential* design alternatives for an actual trial that compares stroke rehabilitation at an IRF to a SNF

Protocol Component	Emulated Trials	Alternative Considerations
Eligibility Criteria		
Patient level	Patients: All Medicare community dwelling fee-for-service acute stroke patients discharged to an IRF or SNF from 2012-2013 Patient level Exclusions: *Patient:* acute LOS>14 days, inpatient stroke, elective admission, metastatic cancer, received care at US territory, < 12 months of continuous Medicare enrollment	Patients must be identified as being in clinical equipoise Patient level Exclusions: *Patient:* acute LOS>14 days, inpatient stroke, elective admission, metastatic cancer, received care at US territory, < 12 months of continuous Medicare enrollment
Hospital level	Hospital level exclusions: Outlier hospital*, discharged < 50 included stroke patients, was not part of a regular use referral triad*	Hospital level exclusions: Outlier hospital*, discharged < 50 included stroke patients, was not part of a regular use referral triad*
Rehabilitation facilities	Rehabilitation level exclusions: **Trial 1:** All rehabilitation facilities used by included hospitals **Trial 2:** Rehabilitation facilities that treated fewer than 5 patients discharged from a single hospital **Trial 3:** Rehabilitation facilities that treated fewer than 10 patients discharged from a single hospital	Rehabilitation level exclusions: Rehabilitation facilities that treated fewer than 10 patients discharged from a single hospital (i.e., Trial 3)
Treatment	IRF versus SNF stroke rehabilitation	IRF versus SNF stroke rehabilitation
Assignment Procedures	Randomization is emulated via 1:1 propensity score matching: Method=Greedy nearest neighbor, caliper=0.1, and match with replacement	Randomization stratified by hospital

Table 5.1 (cont'd)

Follow-up Period	1 year following discharge from acute hospital care (t_0)	3- or 6-months following discharge from acute hospital care (t_0)
Outcome:		
Primary	1-year successful community discharge* (yes or no)	Activity Measure of Post-Acute Care, PROMIS, 16-item Stroke Impact Scale
Secondary:	a) 1-year all-cause mortality b) Time to successful community discharge c) Time to mortality	a) 1-year all-cause mortality b) Time to successful community discharge c) Time to mortality
Causal Contrast	Intention to Treat	Intention to treat
Trial type	Equivalence trial	Non-inferior or superior trial
Analysis plan	d) Risk difference, relative risks, and odds ratios for binary outcomes e) Kaplan-Meier curves and Cox proportional hazard models for time-to-event outcomes f) Fit a local polynomial regression between the matched pair difference over the propensity score to assess for HTE.[40]	a) Risk difference, relative risks, and odds ratios for binary outcomes b) Kaplan-Meier curves and Cox proportional hazard models for time-to-event outcomes
Sensitivity analysis	Competing risks analysis for successful community discharge with death as the competing risk.	Competing risks analysis for successful community discharge with death as the competing risk.

Successful community discharge: discharge home and remained alive and outside of the hospital IRF or SNF care for at least 30 days

Abbreviations: AM-PAC Activity Measure of Post-Acute Care, PROMIS: Patient-Reported Outcomes Measurement Information System, HTE: Heterogeneity of treatment effect

UNIQUE CONTRIBUTIONS OF THIS DISSERTATION

Several elements of this dissertation provide a unique contribution to the literature. Generally, the design of any large RCT requires numerous careful considerations to ensure that the trial is designed to efficiently address the specific question of most interest to patients, clinicians, and policy makers.[153] Often trial design decisions are informed by a heterogenous mix of previous literature, small pilot studies, or simulating results from available datasets. Overall, a novel and unique contribution of this study, was that under an emulated trial framework, we were able to use a single large national database to *both* design and emulate our desired RCT. By linking the analysis to the actual idealized trial design *and* within the idealize trial population – our results provide our "best guess" for the results of an actual trial.

In addition, each of our aims was able to build on the existing literature and provided several new and novel findings. In aim one, similar to previous studies we identified several specific hospital-level factors (e.g. for-profit status, affiliated IRF unit) had large associations with IRF (vs. SNF) discharge.[15–17] However, we believe we are the first study to show that there was substantial variation (i.e., wide 80% Interval Odds Ratios) around the specific contextual effects of these factors. In addition, previous studies have demonstrated large hospital-level variation in IRF (vs. SNF) discharge, but we were the first study to quantify the magnitude of the heterogeneity of hospital effects for individual patients by showing that half of all Medicare stroke patients attended a hospital that changed their probability of being discharged to an IRF (vs. SNF) by over 10%.[15–17] In aim 2, several studies have identified referral networks independently for IRFs or SNFs.[100,109] However, this was the first study to identify and compare these networks specifically for stroke patients. Several previous studies have conducted comparative effectiveness estimates for stroke rehabilitation at IRFs vs. SNFs.[10,37,97,120,121] However, in aim 3, this was the first study to conduct these estimates within a carefully selected target trial population *and* to demonstrate that the relative effect size was conditioned on the specific type of SNF facilities that were included and compared. Additionally, while speculative, our heterogeneity of treatment effect analysis suggests that there may be negative selection of IRF care, because the patients who were least likely to be treated at an IRF, experienced the largest treatment benefit.[137]

LIMITATIONS

There are several important limitations which should be considered when interpreting our results. First, an unbiased propensity score analysis assumes no unmeasured confounders, but Medicare claims data has a limited data range and several important unmeasured confounders likely remain (e.g., baseline function, social support, home environment).[8,19,28,144] The E-values showed that the observed differences in successful community discharge would be nullified if one of these confounders had a RR of 1.64 (Trial 3) to 2.04 (Trial 1). Second**,** claims data is unable to provide the same level of granularity on patient acuity or physiological function that medical records may provide.[89] For this study we had to rely on health service use proxies such as LOS, intensive care unit use, and medical cost quartiles to estimate overall acuity. In addition, we adjusted for the total number of comorbidities and did not have detailed information on the relative severity of these comorbidities. Third, claims data is prone to systematic coding biases and inaccuracies.[154,155] For example, there are known differences in hospital level coding for high revenue invasive procedures (e.g., MRI) compared to low revenue procedures (e.g. electrocardiogram) where there is no financial incentive to report them.[154,155] Fourth**,** in study three, our primary outcome measure of successful community discharge may have biased the results towards favoring IRF patients because potential for community discharge is a clinical indication for IRF referral.[4] However, despite this limitations successful community discharge was developed by CMS as a proxy for a common patient centric measure of function to be used across all types of rehabilitation settings because other activity level functional measures

(Inpatient Rehabilitation Facility – Patient Assessment Instrument and Minimum Data Set) were previously not directly comparable.[127,10] Finally, our data may not be generalizable to stroke patients outside of Medicare and it would be important to replicate this study by using other large databases (e.g., the Veterans Affairs system or private healthcare insurance consortiums).

FUTURE DIRECTIONS

Overall, the body of work in this dissertation provided important information needed to inform the design an RCT to compare stroke rehabilitation at IRFs vs. SNFs. Our work was purely quantitative and relied on retrospective data. Future studies to further inform trial design will need primary qualitative and quantitative data to further enrich design decisions.[156] Several studies that should be considered include: First, a qualitative or mixed methods study among researchers, clinicians and policymakers to gauge interest in the need for such a study, and would provide more in-depth information on the specific comparison (i.e., trial 1, 2, or 3) of most interest. Second, feasibility studies to assess potential patient and hospital facilitators and barriers to recruitment is critical.[101] Third, qualitative studies with clinicians and researchers that help inform trial design considerations (e.g., defining clinical equipoise, patient recruitment, measurement processes, outcome selection, length of follow-up) would further enrich the trial.[156]

Beyond trial design, our results also outline several future lines of work that are needed to more broadly understand stroke rehabilitation at IRFs and SNFs. First, an updated cost effectiveness study that compares both direct and indirect costs of IRF and SNF care is needed. This study should assess costs associated with the initial stay as well as long term total medical costs. In this dissertation, the two cost estimates that were frequently cited came from two studies that used Medicare data either from 2002-2003[157] or from 1997.[10] Second, studies that focus on characterizing the amount of variation in outcomes, as well as drivers of this variation for patients treated at SNFs are needed. In aim three, we showed that the relative effect size for IRF (vs. SNF) rehabilitation was almost entirely driven by the type of SNFs that were compared. A 2017 systematic review of 13 studies by Alcusky et al 2017 evaluated the association of facility level characteristics on outcomes for stroke patients and found that length of stay was the only process of care measure that was ever captured.[27] In addition, most studies focused almost exclusively on IRFs while ignoring SNFs.[9,23,27,107] Third, for comparative effectiveness estimates, alternative methods such as instrumental variables (IV) analyses could be considered. IVs are variables that are highly correlated with treatment selection but are not associated with the outcome. A previous study used both PS and IVs to compare IRF to SNF care.[97] This study found smaller effect sizes for IRF care using differential distance and %IRF discharge as IVs (See table 1.2 in Chapter 1).[97] However, for this dissertation, new IVs would have to be identified as %IRF discharge would not be as strong of an IVs, on account that we pre-selected hospitals with the smallest effect sizes (typical hospitals) and we did not have data to calculate the differential distance from a patients home to the closest IRF or SNF.

CONCLUSION

In sum, there is substantial variation in IRF and SNF use for stroke patients across the United States which has very large implications for healthcare expenditure and patient outcomes. Access of IRF and SNF care is complex, multi-dimensional and poorly understood which limits the ability of observational data to make valid unbiased comparative effectiveness estimates. Thus, an RCT is needed. In this dissertation, we conducted a series of three studies to inform the design of such a trial. First, we showed that for many patients, hospitals were major selection forces that influence whether they are discharged to an IRF or SNF. Second, we carefully selected the hospitals and patients that we believed represented the optimal target trial population. Finally, we emulated the three pragmatic RCTs that we believed represented the desired alternative trial designs and showed that while IRF patients did better, the relative differences were contingent on the type of SNFs that were included. Finally, we outlined several important next steps that are needed to continue with designing such a trial.

APPENDICES

Appendix A: Supplemental Tables

Supplemental Table 2.1: Data sources used to assemble the final cohort of acute Medicare stroke patients who were discharged to receive care at an IRF or SNF

Data Source	Description	Abstracted data	Aims file was used in
Patient level data			
Inpatient claims (IPC) (2011-2014)		Stroke patients, stroke subtype, stroke severity, clinical comorbidities, acute and IRF length of stay	Aims 1,2,3
MedPAR (2011-2014)	Provides highly aggregated data for charges and length of stay for single hospitalizations/SNF stays	Charge data, service use data, and SNF length of stay	Aims 1,2,3
Master Beneficiary Summary File (MBSF) (2011-2014)		Age, race, sex, patient zip code	Aims 1,2,3
Part B Carrier Summary Data File (Part B file) (2012-2013)	Carrier level summary Current Procedure Terminology (CPT) codes as well as information on the number of allowed services, charges, and payments	Number of physical, occupational, and speech language therapy Current Procedural Therapy (CPT) codes provided during the inpatient stay	Aims 1,2,3
Zip code level data			
American Community Survey (ACS) (2013)	Provides census level information for zip code level data from a sample of patients	Zip code level data for median income, and proportion of the population with a bachelor's degree	Aims 1 and 2

Supplemental Table 2.1 (cont'd)

Hospital level data			
Provider of Service (POS) (2012-2013)	File contains detailed information on hospitals which are linked to the final cohort through a unique hospital identifier (PRVDR_NUM)	Hospital and geographic level characteristics	Aims 1 and 2
Hospital compare data (2014)	Hospital compare data provides quality data for Medicare certified hospitals. Process and outcome measures are chosen by CMS, hospital industry, and public sector stakeholders.	Hospital process and outcome data	Aims 1 and 2

Abbreviations:
MedPAR: Medicare Provider Analysis and Review

Supplemental Table 2.2: Technical description of all covariates used to characterize heterogeneity of hospital effects among acute Medicare stroke patients who were discharged to receive care at an IRF or SNF

	Definition	Parameterization	Additional Information	File:
Sociodemographic characteristics				
Age		Per 1-year increase	Age at time of stroke	MBSF
Race		White, Black, Hispanic, Other		MBSF
Sex		Male or Female		MBSF
Median annual household income (per $1,000)		<25k 25-50k 50-75k 75-100k >100k Missing	Race matched to zip code	ACS and MBSF
Proportion of population with a bachelor's degree		<10% 10-15% 15-20% 20-30% 30-45% >45% Missing	Sex and age matched to zip code	ACS and MBSF
Pre-Stroke functional proxies				
Previous home-time	Days alive and spent outside of acute, IRF or SNF care	Per 1-day increase	Taken 1 year prior to the indexed stroke	MedPAR
Previous number of hospitalizations	Number of acute hospitalizations	Per 1-hospitalization increase	Taken 1 year prior to the indexed stroke	IPC
Previous IRF use	Any IRF use	Yes/No	Taken 1 year prior to the indexed stroke	IPC
Previous SNF use	Any SNF use	Yes/No	Taken 1 year prior to the indexed stroke	MedPAR
Social Security Disability	Identified as being disabled by Social Security	Yes/no		MBSF

Supplemental Table 2.2 (cont'd)

Stroke characteristics				
Stroke subtype	Ischemic: ICD-9 (431, 433.x1) Intracerebral hemorrhagic: ICD-9 (434.x1)	Ischemic or intracerebral hemorrhagic	Documented during indexed stroke hospitalization	IPC
Stroke administrative Severity Index[72]		Mild Moderate Severe	Based on five ICD 9 diagnostic codes (aphasia, coma, dysarthria/dysphagia, hemiplegia/monoplegia, and neglect) and 2 ICD-9 procedure codes (nutritional infusion and tracheostomy/ventilation) documented during indexed stroke hospitalization	IPC
Emergency department Admission		Yes/No	Based on any emergency room charge data	MBSF
Elixhauser comorbidities (based on ICD-9 codes documented as present prior to acute admission)				IPC
Valvular disorders	'0932','7463','7464','7465','7466','V422','V433' '394','395','396','397','424'	Yes/No		
Pulmonary circulatory disorders	'4150','4151','4170','4178','4179' '416'	Yes/No		
Peripheral vascular disorders	'0930','4373','4431','4432','4438','4439','4471','5571','5579' ,'V434' '440','441'	Yes/No		
Uncomplicated hypertension	'401'	Yes/No		
Complicated hypertension	'402','403','404','405'	Yes/No		

Supplemental Table 2.2 (cont'd)

Paralysis	'3341','3440','3441','3442','3443','3444','3445','3446','3449' '342','343'	Yes/No
Other neurological disorders	'3319','3320','3321','3334','3335','3362','3481','3483','7803','7843' '33392' '334','335','340','341','345'	Yes/No
Chronic pulmonary disorders	'4168','4169','5064','5081','5088' '490','491','492','493','494','495','496','500','501','502','503','504','505'	Yes/No
Uncomplicated diabetes	'2500','2501','2502','2503'	Yes/No
Complicated diabetes	'2504','2505','2506','2507','2508','2509'	Yes/No
Hypothyroidism	'2409','2461','2468', '243','244'	Yes/No
Renal failure	'5880','V420','V451', '40301','40311','40391','40402','40403','40412','40413','40492','40493','585','586','V56'	Yes/No
Liver disease	'0706','0709','4560','4561', '4562','5722','5723','5724','5728','5733','5734','5738','5739','V427', '07022','07023','07032','07033','07044','07054', '570','571'	Yes/No
Peptic ulcer disease excluding bleeding	'5317','5319','5327','5329','5337','5339','5347','5349'	Yes/No
AIDS/HIV	'042','043','044'	Yes/No

Supplemental Table 2.2 (cont'd)

Lymphoma	'2030','2386', '200','201','202'	Yes/No
Metastatic cancer	'196','197','198','199'	Yes/No
Solid tumor without metastasis	'140','141','142','143','144','145','146','147','148','149','150','151','152', '153','154','155','156','157','158','159','160','161','162','163','164','165','166','167', '168','169','170','171','172','174','175','176','177','178','179','180','181','182','183', '184','185','186','187','188','189','190','191','192','193','194','195'	Yes/No
Rheumatoid arthritis/collagen	'7010','7100','7101','7102','7103','7104','7108','7109','7112','7193','7285', '446','714','720','725'	Yes/No
Coagulopathy	'2871','2873','2874','2875', '286'	Yes/No
Obesity	'2780'	Yes/No
Weight Loss	'7832','7994', '260','261','262','263'	Yes/No
Fluid and electrolyte Disorders	'2536', '276'	Yes/No
Blood loss anemia	'2800'	Yes/No
Iron deficiency anemia	'2801','2808','2809', '281'	Yes/No

Supplemental Table 2.2 (cont’d)

Alcohol abuse	'2652','2911','2912','2913','2915','2918','2919','3030','3039','3050','3575','4255','5353','5710','5711','5712','5713','V113', '980'	Yes/No		
Drug abuse	'3052','3053','3054','3055','3056','3057','3058','3059', '292','304', 'V6542'	Yes/No		
Psychoses	'2938', '295','297','298', '29604','29614','29644','29654'	Yes/No		
Depression	'2962','2963','2965','3004', '309','311'	Yes/No		
Total Elixhauser comorbidity index		Per-1 comorbidity increase (range 0-31)	Total number of comorbidities	
Comorbidities not included in the Elixhauser index				
Dementia	'2900','2901','2902','2903','2904','2912','2941','3310','3311','3312', ‘29010','29011','29012','29013','29020','29021','29040','29041','29042','29043', '29410','33111', '33119', '33182'	Yes/No	ICD-9 codes documented as present prior to acute admission	IPC
End Stage Renal Disease		Yes/No	Medicare Enrollment Reason	MBSF
Stroke Symptoms (based on ICD-9 codes which were *not* present prior to acute admission				IPC
Coma	'78001', '78003'	Yes/No		
Aphasia	7843', '43811'	Yes/No		
Dysphagia/dysarthria	‘7872’ '43813', '43882', '78451'	Yes/No		

Supplemental Table 2.2 (cont'd)

Hemiplegia/ monoplegia	'4382', '4383','4384','4385' '34290'	Yes/No	
Neglect	'7818'	Yes/No	
Life Saving Procedures (ICD-9 procedures documented as occurring during the indexed hospitalization)			
Hemodialysis	'3995'	Yes/No	IPC
Gastrostomy tube	'4311', '4319', '4432', '432'	Yes/No	
Intubation/ventilation	'9604', '9605', '9671', '9672', '9673', '9674', '9675', '9676', '9677', '9678','9678' '9679'	Yes/No	
CPR	'9960', '9963'	Yes/No	
Parenteral nutrition	'9915', '966'	Yes/No	
tPA	'9910'	Yes/No	
Hospital health services use			
Length of Stay	Length of stay during the acute hospitalization	Per 1-day increase	IPC
ICU days	Number of days in an ICU unit	Per 1-day increase	MedPAR
CCU days	Number of days in a CCU unit	Per 1-day increase	MedPAR
In hospital rehabilitation services			
Total Number of PT "encounters"	CPT codes: 0420, 0421, 0422, 0423, 0424, 0429	Per increase in 1 CPT code	
Total Number of OT "encounters"	CPT codes: 0430, 0431, 0432, 0433, 0434, 0439		
Total Number of SLT "encounters"	CPT codes: 0440, 0441, 0442, 0443, 0444, 0449		
Hospital charge data (Total charges during the indexed acute hospitalization)			MedPAR
Total		Quartiles 1-4	
Pharmacy		Quartiles 1-4	
Labs		Quartiles 1-4	
Radiology		Quartiles 1-4	

Supplemental Table 2.2 (cont'd)

In hospital service use				MedPAR
Inhalation services				
MRI services		Yes/No		
Operating room services		Yes/No		
		Hospital Characteristics:		POS
Bed Size		Per 50 bed increase		
Residency program		Yes/No	Combination measures of having either an allopathic or osteopathic residency program	
Medical School Affiliation		Yes/No		
Hospital ownership		Church Private-not for profit Private-for profit Government Other		
Hospital affiliated IRF unit		Yes/No		
Hospital has approved swing beds	Swing beds=acute care hospital can provide SNF care if beds are available	Yes/No		

Supplemental Table 2.2 (cont'd)

Hospital process scores			CPT
Composite hospital process sum score	Per 1-point increase	Step 1: assign points to each hospital based on percentage of patients (missing data=0 points, <90%=1 point, 90-94%=2 points, 95-99%=3 points, and 100%=4 points) for eight stroke services (venous thrombosis prophylaxis, anti-thrombotic use, anti-coagulation uses for atrial fibrillation/flutter, anti-thrombotic use, anti-thrombotic use on day two, discharged on a statin, stroke education, stroke rehabilitation assessment). Step 2: Sum scores from eight measures	

Supplemental Table 2.2 (cont'd)

Hospital outcome score			CPT
Total hospital outcome score	*<u>Better than national average</u>* in either 30-day all-cause mortality and/or 30-day all-cause readmissions *<u>No different from national average</u>* in both 30-day all-cause mortality and 30-day all-cause readmissions *<u>Worse than national average</u>* in either 30-day all-cause mortality and/or 30-day all-cause readmissions *<u>Missing</u>* or not enough information	Combined 30-day all-cause mortality and 30-day all-cause readmissions. Scores are nationally adjusted measures.	
	Geographic characteristics		
Hospital setting	Urban or rural		POS
CMS region	CMS regions 1-10		POS
1	CT, ME, MA, NH, RI, VT		
2	NY, NJ		
3	DE, DC, MD, PA, VA, WV		
4	AL, FL, GA, KY, MS, NC, SC, TN		
5	IL, IN, MI, MN, OH, WI		
6	AR, LA, NM, OK, TX		
7	IA, KS, MO, NE		
8	CO, MT, ND, SD, UT, WY		
9	AZ, CA, HI, NV		
10	AK, ID, OR, WA		

Files: IPC: Inpatient claims, MedPAR Medicare Provider Analysis and Review, MBSF: Master Beneficiary Summary File, ACS: American Community Survey, POS: Provider of Service, CPT: Current Procedural Terminology

Supplemental Table 2.3: All unadjusted and adjusted odds ratio associations of selected patient and hospital contextual factors with IRF (vs. SNF) discharge among Medicare stroke survivors-multivariable logistic regression results

	Model 1 (OR)	95% CI	Model 2 (aOR)	95% CI	Model 3 (aOR)	95% CI
Sociodemographic characteristics						
Age	0.942	[0.940,0.943]	0.932	[0.930,0.933]	0.932	[0.930,0.933]
Race						
Black	0.942	[0.903,0.983]	0.903	[0.858,0.951]	0.897	[0.852,0.944]
Hispanic	**1.131**	[1.067,1.199]	0.975	[0.907,1.048]	0.957	[0.891,1.029]
Other	0.977	[0.915,1.044]	1.035	[0.957,1.119]	1.029	[0.951,1.113]
Female sex	0.768	[0.748,0.788]	0.735	[0.713,0.757]	0.735	[0.713,0.757]
Median annual household income (based on ZIP code aggregate data)						
< 25k	1.017	[0.947,1.093]	0.957	[0.881,1.039]	0.962	[0.886,1.045]
25-50k	0.924	[0.898,0.951]	0.924	[0.892,0.957]	0.929	[0.897,0.962]
75-100k	1.002	[0.964,1.041]	1.032	[0.986,1.081]	1.035	[0.989,1.084]
>100k	1.015	[0.961,1.072]	1.075	[1.005,1.151]	1.078	[1.007,1.154]
Missing	1.039	[0.716,1.508]	0.981	[0.646,1.490]	0.992	[0.654,1.507]
Proportion of population with bachelor’s degree (based on ZIP code aggregate data)						
<10%	1.065	[1.023,1.109]	1.002	[0.956,1.050]	1.001	[0.955,1.049]
10-15%	1.026	[0.987,1.068]	1.004	[0.960,1.050]	1.004	[0.960,1.050]
20-30%	0.982	[0.945,1.021]	0.989	[0.947,1.034]	0.990	[0.947,1.034]
30-45%	1.010	[0.967,1.055]	1.025	[0.975,1.078]	1.025	[0.975,1.078]
>45%	1.119	[1.059,1.182]	1.110	[1.041,1.184]	1.108	[1.039,1.182]
Missing	0.935	[0.642,1.361]	0.966	[0.634,1.473]	0.957	[0.628,1.459]
Social security disability	0.935	[0.830,1.054]	0.896	[0.784,1.024]	0.897	[0.785,1.025]
Pre-stroke functional proxies (Taken 1 year prior to the indexed stroke)						
Pre-home-time	1.008	[1.007,1.009]	1.009	[1.008,1.011]	1.009	[1.008,1.011]
Previous number of hospitalizations	0.923	[0.904,0.942]	0.923	[0.903,0.944]	0.924	[0.903,0.944]
Previous IRF use	2.847	[2.629,3.084]	1.884	[1.729,2.052]	1.881	[1.727,2.050]
Previous SNF use	0.393	[0.368,0.419]	0.391	[0.364,0.419]	0.391	[0.365,0.419]

Supplemental Table 2.3 (cont'd)

Comorbidities						
Elixhauser comorbidity index	0.978	[0.944,1.013]	0.978	[0.940,1.017]	0.978	[0.940,1.017]
CHF	1.028	[0.980,1.078]	1.019	[0.966,1.074]	1.018	[0.965,1.073]
Arrhythmia	1.045	[1.001,1.090]	1.048	[0.999,1.098]	1.048	[1.000,1.099]
Valvular disease	1.032	[0.982,1.085]	1.036	[0.981,1.095]	1.037	[0.982,1.096]
Pulmonary circulatory disorder	1.010	[0.940,1.084]	1.008	[0.932,1.090]	1.009	[0.933,1.091]
Peripheral vascular diseases	0.998	[0.947,1.052]	0.995	[0.939,1.054]	0.995	[0.940,1.055]
Uncomplicated hypertension	1.053	[1.001,1.107]	1.046	[0.989,1.106]	1.045	[0.988,1.105]
Complicated hypertension	1.071	[0.990,1.158]	1.090	[0.999,1.188]	1.089	[0.998,1.187]
Paralysis	1.059	[1.014,1.106]	1.055	[1.006,1.107]	1.054	[1.005,1.106]
Other neurological condition	1.027	[0.982,1.073]	1.022	[0.973,1.073]	1.022	[0.974,1.073]
COPD	1.026	[0.976,1.078]	1.038	[0.982,1.097]	1.039	[0.984,1.098]
Uncomplicated diabetes	1.006	[0.963,1.052]	1.006	[0.958,1.056]	1.005	[0.957,1.056]
Complicated diabetes	1.026	[0.961,1.095]	1.032	[0.960,1.109]	1.032	[0.960,1.109]
Hypothyroidism	1.031	[0.983,1.080]	1.016	[0.964,1.070]	1.016	[0.964,1.070]
Renal failure	1.008	[0.933,1.088]	0.986	[0.906,1.074]	0.987	[0.907,1.074]
Liver disease	1.054	[0.931,1.192]	1.042	[0.908,1.196]	1.041	[0.907,1.195]
Peptic ulcer w/o bleed	1.068	[0.938,1.217]	1.065	[0.922,1.230]	1.064	[0.921,1.229]
HIV	1.030	[0.465,2.282]	0.923	[0.404,2.110]	0.914	[0.399,2.094]
Lymphoma	1.083	[0.906,1.296]	1.195	[0.980,1.457]	1.194	[0.979,1.456]
Sold tumor w/o metastasis	1.059	[0.963,1.163]	1.086	[0.979,1.205]	1.087	[0.980,1.206]

Supplemental Table 2.3 (cont'd)

Rheumatoid arthritis	1.023	[0.950,1.100]	1.032	[0.952,1.119]	1.033	[0.953,1.121]
Coagulopathy	0.963	[0.894,1.037]	0.975	[0.898,1.058]	0.975	[0.898,1.059]
Obesity	1.026	[0.968,1.088]	1.038	[0.973,1.108]	1.039	[0.973,1.108]
Weight loss	0.994	[0.926,1.066]	0.967	[0.895,1.046]	0.964	[0.892,1.043]
Fluid electrolyte disorder	1.016	[0.970,1.064]	1.033	[0.982,1.087]	1.032	[0.981,1.086]
Blood loss anemia	0.962	[0.789,1.174]	0.964	[0.774,1.202]	0.963	[0.773,1.200]
Iron deficient anemia	1.005	[0.928,1.088]	1.043	[0.955,1.138]	1.042	[0.954,1.137]
Alcohol abuse	1.016	[0.735,1.404]	1.021	[0.712,1.464]	1.021	[0.712,1.463]
Drug abuse	1.127	[0.726,1.748]	1.217	[0.751,1.970]	1.219	[0.753,1.972]
Psychosis	1.014	[0.918,1.119]	0.975	[0.874,1.087]	0.975	[0.874,1.087]
Dementia	0.349	[0.334,0.366]	0.299	[0.284,0.314]	0.298	[0.284,0.314]
ESRD	0.752	[0.591,0.957]	0.722	[0.553,0.944]	0.720	[0.551,0.941]
Stroke characteristics:						
Stroke subtype (ref=ischemic)						
ICH	1.001	[0.961,1.044]	1.002	[0.957,1.049]	1.001	[0.956,1.049]
Stroke severity (ref=mild)						
Moderate	1.007	[0.961,1.055]	1.004	[0.954,1.057]	1.005	[0.955,1.058]
Severe	1.011	[0.931,1.099]	0.999	[0.911,1.095]	1.000	[0.912,1.096]
Stroke symptoms						
Aphasia	0.993	[0.956,1.031]	0.982	[0.942,1.025]	0.983	[0.942,1.025]
Coma	1.094	[0.854,1.402]	1.043	[0.795,1.369]	1.041	[0.793,1.366]
Dysarthphagia	1.000	[0.961,1.040]	1.003	[0.961,1.048]	1.002	[0.960,1.047]
Hemimonoplegia	1.013	[0.967,1.062]	1.010	[0.958,1.064]	1.009	[0.957,1.063]
Neglect	0.986	[0.901,1.078]	1.009	[0.914,1.114]	1.010	[0.914,1.115]
Hospital Health Services Use						
LOS (1-day Inc.)	0.987	[0.982,0.993]	0.993	[0.987,0.999]	0.993	[0.987,0.999]

Supplemental Table 2.3 (cont'd)

Intensive care days used by beneficiary for stay	0.997	[0.992,1.002]	0.990	[0.983,0.997]	0.990	[0.983,0.997]
Coronary care days used by beneficiary for stay	0.975	[0.968,0.982]	0.960	[0.950,0.970]	0.961	[0.951,0.972]
ED admission	0.855	[0.821,0.890]	0.967	[0.922,1.015]	0.973	[0.928,1.021]
Lifesaving procedures						
Hemodialysis	0.711	[0.637,0.793]	0.668	[0.592,0.753]	0.668	[0.593,0.754]
GI tube	0.468	[0.441,0.497]	0.419	[0.393,0.448]	0.420	[0.393,0.449]
CPR	1.132	[0.658,1.948]	1.339	[0.742,2.417]	1.345	[0.746,2.426]
Parenteral nutrition	0.973	[0.901,1.050]	1.084	[0.996,1.180]	1.091	[1.002,1.187]
Intubation/ventilation	1.044	[0.953,1.145]	1.198	[1.083,1.325]	1.201	[1.085,1.329]
tPA	1.817	[1.723,1.917]	2.092	[1.970,2.222]	2.096	[1.974,2.226]
Stroke symptoms						
Aphasia	0.993	[0.956,1.031]	0.982	[0.942,1.025]	0.983	[0.942,1.025]
Coma	1.094	[0.854,1.402]	1.043	[0.795,1.369]	1.041	[0.793,1.366]
Dysarthphagia	1.000	[0.961,1.040]	1.003	[0.961,1.048]	1.002	[0.960,1.047]
Hemimonoplegia	1.013	[0.967,1.062]	1.010	[0.958,1.064]	1.009	[0.957,1.063]
Neglect	0.986	[0.901,1.078]	1.009	[0.914,1.114]	1.010	[0.914,1.115]
Number of PT CPT revenue codes (ref=0)						
1-3	1.548	[1.416,1.692]	1.845	[1.652,2.062]	1.857	[1.662,2.074]
4-7	1.462	[1.336,1.600]	1.924	[1.720,2.152]	1.940	[1.735,2.170]
8-11	1.310	[1.192,1.439]	1.888	[1.680,2.122]	1.908	[1.698,2.144]
>11	1.397	[1.266,1.542]	2.092	[1.849,2.366]	2.118	[1.873,2.396]

Supplemental Table 2.3 (cont'd)

Number of OT CPT revenue codes (ref=0)						
1-2	1.684	[1.625,1.746]	1.702	[1.626,1.782]	1.684	[1.609,1.763]
3-6	1.799	[1.735,1.865]	1.990	[1.899,2.086]	1.967	[1.877,2.062]
7-9	1.779	[1.692,1.871]	2.053	[1.928,2.185]	2.027	[1.904,2.158]
>9	1.755	[1.652,1.864]	2.151	[1.996,2.318]	2.123	[1.970,2.287]
Number of SLT CPT revenue codes (ref=0)						
1-2	1.256	[1.216,1.297]	1.296	[1.249,1.345]	1.293	[1.246,1.342]
3-5	1.181	[1.141,1.222]	1.268	[1.219,1.320]	1.266	[1.216,1.317]
6-7	1.128	[1.070,1.190]	1.233	[1.161,1.310]	1.232	[1.160,1.308]
>7	1.129	[1.065,1.196]	1.255	[1.173,1.343]	1.252	[1.170,1.340]
Charge data						
Total charge quartiles (ref=quartile 1)						
Quartile 2	1.042	[1.008,1.077]	1.022	[0.985,1.061]	1.021	[0.984,1.059]
Quartile 3	1.067	[1.031,1.105]	1.034	[0.995,1.075]	1.032	[0.992,1.073]
Quartile 4	1.087	[1.043,1.132]	1.035	[0.988,1.083]	1.031	[0.984,1.079]
Pharmacology charge quartiles (ref=quartile 1)						
Quartile 2	0.837	[0.809,0.866]	0.787	[0.756,0.819]	0.787	[0.756,0.818]
Quartile 3	0.718	[0.692,0.744]	0.639	[0.611,0.668]	0.638	[0.610,0.667]
Quartile 4	0.641	[0.614,0.669]	0.529	[0.501,0.559]	0.527	[0.499,0.556]
Laboratory charge quartiles (ref=quartile 1)						
Quartile 2	1.012	[0.978,1.048]	0.793	[0.761,0.827]	0.789	[0.757,0.822]
Quartile 3	0.972	[0.937,1.008]	0.667	[0.637,0.700]	0.660	[0.629,0.692]
Quartile 4	0.900	[0.863,0.939]	0.530	[0.500,0.563]	0.519	[0.490,0.551]
Radiology charge quartiles (ref=quartile 1)						
Quartile 2	1.127	[1.088,1.166]	1.006	[0.965,1.048]	1.004	[0.964,1.046]
Quartile 3	1.223	[1.180,1.268]	1.106	[1.059,1.155]	1.100	[1.054,1.149]
Quartile 4	1.486	[1.428,1.546]	1.335	[1.270,1.402]	1.322	[1.258,1.389]

Supplemental Table 2.3 (cont'd)

Hospital service use (yes/no)						
Inhalation therapy	0.900	[0.878,0.923]	0.902	[0.874,0.931]	0.905	[0.877,0.934]
MRI	1.335	[1.301,1.371]	1.385	[1.344,1.428]	1.382	[1.341,1.425]
Operating room	1.030	[0.989,1.072]	1.105	[1.056,1.157]	1.108	[1.059,1.160]
Hospital Characteristics						
Hospital bed count (per 50 increase)					1.000	[0.987,1.013]
Hospital process sum score					0.999	[0.987,1.011]
Hospital outcome score (ref=national average)						
Below Average					0.885	[0.707,1.107]
Above Average					1.119	[0.878,1.425]
Missing					0.320	[0.124,0.825]
Hospital ownership (ref= Private not for-profit						
Church					1.092	[0.924,1.291]
Private-for profit					1.339	[1.136,1.578]
Government					0.808	[0.643,1.016]
Other					0.968	[0.837,1.118]
Medical school affiliation					1.173	[1.018,1.351]
Residency					0.940	[0.809,1.092]
IRF unit					2.527	[2.251,2.837]
Swing bed					0.790	[0.571,1.094]

Supplemental Table 2.3 (cont'd)

CMS region (Ref=5 (IL, IN, MI, MN, OH, WI)		
1) CT, ME, MA, NH, RI, VT	1.052	[0.813,1.360]
2) NY, NJ	1.388	[1.111,1.734]
3) DE, DC, MD, PA, VA, WV	1.244	[1.010,1.531]
4) AL, FL, GA, KY, MS, NC, SC, TN	1.300	[1.090,1.551]
6) AR, LA, NM, OK, TX	3.120	[2.545,3.825]
7) IA, KS, MO, NE	1.689	[1.299,2.197]
8) CO, MT, ND, SD, UT, WY	0.990	[0.709,1.383]
9) AZ, CA, HI, NV	1.351	[1.103,1.655]
10) AK, ID, OR, WA	0.675	[0.497,0.917]
Urban (vs. rural)	1.710	[1.444,2.026]

Abbreviations: IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility, ED: Emergency department, ICU: Intensive care unit, CCU: Cardiology care unit, tPA: Tissue plasminogen activator, CMS: Center for Medicare & Medicaid Services, OR: Odds Ratio, aOR: Adjusted Odds Ratio

Model 1: Single level logistic regression model that modeled discharge to an IRF vs. SNF that included patient level fixed effects

Model 2: Multi-level logistic regression model that modeled discharge to an IRF vs. SNF that included patient level fixed effects and a hospital random effect

Model 3: Multi-level logistic regression model that modeled discharge to an IRF vs. SNF that included patient and hospital level fixed effects and a hospital random effect

Supplemental Table 3.1: Baseline hospital characteristics presented at the patient level for Medicare stroke patients who were discharge from 3,039 hospitals to an IRF or a SNF

	(n=145,894 patients) (%)
Number of hospitals	3,039
Number of beds (SD)	437.9 (325.7)
Total hospital process sum score (SD)	15.1 (4.9)
Combined mortality and rehospitalizations outcome score	
Worse than national average	9.2
National Average	81.8
Better than national average	7.4
Missing	1.5
Hospital ownership	
Church	14.0
Private not for profit	46.7
Private for profit	13.2
Government	5.8
Other	20.3
IRF affiliated unit	56.6
Medical school affiliation	50.5
Urban hospital	88.4
CMS region	
1) CT, ME, MA, NH, RI, VT	5.8
2) NY, NJ	9.5
3) DE, DC, MD, PA, VA, WV	11.3
4) AL, FL, GA, KY, MS, NC, SC, TN	22.4
5) IL, IN, MI, MN, OH, WI	18.4
6) AR, LA, NM, OK, TX	11.7
7) IA, KS, MO, NE	5.2
8) CO, MT, ND, SD, UT, WY	2.3
9) AZ, CA, HI, NV	10.3
10) AK, ID, OR, WA	3.3

Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, SD: Standard deviation, CMS: Centers for Medicaid and Medicare Services

Total hospital process sum score: Combined score for proportion of patients that received eight stroke quantity process measures

Supplemental Table 3.2: Adjusted associations of patient and hospital level factor associations with IRF (vs. SNF) discharge among Medicare acute stroke patients discharged to IRF or SNF (i.e. cases) identified from the multi-level logistic regression model

	aOR	95 % CI	*p*-value
Sociodemographic			
Age	0.932	[0.930,0.933]	<0.01
Race (ref=white)			
Black	0.896	[0.852,0.944]	<0.01
Hispanic	0.974	[0.907,1.047]	0.48
Other	1.035	[0.957,1.119]	0.39
Female sex	0.720	[0.701,0.740]	<0.01
Median annual income (ref=50-75k)			
< 25k	0.955	[0.880,1.035]	0.26
25-50k	0.917	[0.887,0.949]	<0.01
75-100k	1.045	[0.999,1.093]	0.05
>100k	1.125	[1.056,1.199]	<0.01
Missing	0.941	[0.853,1.039]	0.23
Prior health care utilization*			
Previous home-time	1.009	[1.008,1.011]	<0.01
Previous number of hospitalizations	0.922	[0.901,0.942]	<0.01
Previous IRF use	1.876	[1.722,2.043]	<0.01
Previous SNF use	0.390	[0.364,0.418]	<0.01
Comorbidities			
Total Elixhauser score	1.003	[0.996,1.010]	0.45
Dementia	0.298	[0.284,0.313]	<0.01
Stroke Characteristics			
Stroke subtype (Ref=Ischemic)			
Intracerebral hemorrhagic	0.997	[0.953,1.043]	0.90
Stroke severity* (ref=Mild)			
Moderate	1.005	[0.976,1.034]	0.75
Severe	1.002	[0.969,1.038]	0.89
Hospital Service use			
Length of stay	0.994	[0.989,0.999]	0.023
ICU use	1.202	[1.161,1.244]	<0.01
ED admission	1.006	[0.960,1.055]	0.79
Lifesaving procedures			
Hemodialysis	0.628	[0.559,0.705]	<0.01
Gastrostomy tube	0.413	[0.387,0.442]	<0.01
CPR	1.297	[0.720,2.335]	0.39
Parenteral nutrition	1.082	[0.994,1.178]	0.07
Intubation/ventilation	1.168	[1.057,1.292]	<0.01
tPA	2.102	[1.980,2.232]	<0.01

Supplemental Table 3.2 (cont'd)

Number of physical therapy CPT revenue codes (ref=0)			
1-3	1.828	[1.637,2.043]	<0.01
4-7	1.900	[1.699,2.125]	<0.01
8-11	1.865	[1.660,2.095]	<0.01
>11	2.050	[1.813,2.318]	<0.01
Number of occupational therapy CPT revenue codes (ref=0)			
		[1.623,1.778]	<0.01
3-6	1.980	[1.889,2.075]	<0.01
7-9	2.046	[1.923,2.178]	<0.01
>9	2.135	[1.982,2.300]	<0.01
Number of speech language therapy CPT revenue codes (ref=0)			
1-2	1.288	[1.241,1.337]	<0.01
3-5	1.257	[1.208,1.308]	<0.01
6-7	1.219	[1.148,1.294]	<0.01
>7	1.241	[1.160,1.328]	<0.01
Pharmacy charges (ref=quartile 1)			
Quartile 2	0.790	[0.760,0.822]	<0.01
Quartile 3	0.645	[0.617,0.674]	<0.01
Quartile 4	0.530	[0.502,0.559]	<0.01
Laboratory charges (ref=quartile 1)			
Quartile 2	0.790	[0.758,0.823]	<0.01
Quartile 3	0.664	[0.634,0.696]	<0.01
Quartile 4	0.531	[0.501,0.563]	<0.01
Hospital Services use (yes/no)			
Inhalation therapy	0.900	[0.872,0.929]	<0.01
MRI	1.351	[1.311,1.391]	<0.01
Operating room	1.126	[1.077,1.178]	<0.01
var(_cons[prvdr_num])	5.203	[4.576,5.915]	

Median annual household income: taken from race matched zip code data Prior health care utilization Taken 1 year prior to the indexed stroke event. Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator

Supplemental Table 3.3: Number and type of hospital referral networks for discharging acute stroke patients (i.e. cases) to receive IRF or SNF care used by each potential trial sample

	All Hospitals		**Typical Hospitals**		
Included hospitals (n= hospitals)	All (n=3,039)	≥20 cases (n=1,816)	≥20 cases (n=891)	≥50 cases (n=475)	≥100 cases (n=169)
Acute care hospitals (n= hospitals)	145,984	135,415	60,529	47,326	25,980
% IRF Discharge (SD) [range]	0.48 (0.21) [0-1.00]	0.49 (0.20) [0-1.00]	0.47 (0.12) [0-0.76]	0.48 (0.10) [0.17-0.74]	0.49 (0.08) [0.27-0.64]
Minimum use referral network (i.e. single hospital discharges >1 patient to a specific rehabilitation facility)					
IRF referral networks					
Number of hospitals with any IRF use	2,542	1,769	890	475	169
Number of IRFs	1,150	1,133	950	782	545
Mean Number of cases at IRFs (SD):	60.61 (56.58)	58.74 (56.46)	23.73 (35.73)	28.95 (35.81)	23.21 (34.62)
Mean number of IRFs with any use by each hospital (SD):	2.60 (2.53)	3.11 (2.85)	2.93 (2.57)	3.62 (3.12)	5.20 (4.03)
SNF referral networks					
Number of hospitals with any SNF use	2,981	1,815	891	475	169
Number of SNFs	12,401	11,772	7,855	6,352	3,932
Mean number of SNF cases (SD):	6.15 (6.17)	5.85 (6.16)	4.11 (4.49)	3.89 (4.44)	3.39 (4.07)
Mean number of SNFs with any use by each hospital (SD):	9.50 (9.76)	13.58 (10.56)	13.51 (10.29)	18.72 (11.45)	28.32 (13.21)

Supplemental Table 3.3 (cont'd)

Regular use referral network (i.e. single hospital discharges >5 patients to a specific rehabilitation facility)					
IRF referral networks					
Number of hospitals with at least 1 regularly used IRF	1,764	1,546	823	475	169
Number of IRFs regularly used	1,103	1,046	658	480	256
Mean number of "regularly used" IRFs by each hospital (SD):	1.33 (0.82)	1.38 (0.87)	1.29 (0.69)	1.44 (0.85)	1.79 (1.15)
SNF referral networks					
Number of hospitals with at least 1 regularly used SNF	1,736	1,337	725	441	166
Number of SNFs regularly used	3,773	3,492	1,737	1,338	712
Mean number of "regularly used" SNFs by each hospital (SD):	2.43 (1.98)	1.17 (0.51)	2.55 (1.89)	3.22 (2.10)	4.49 (2.60)
Frequent use referral network (i.e. single hospital discharges >10 patients to a specific rehabilitation facility)					
IRF referral networks					
Number of hospitals with at least 1 frequently used IRF	1,397	1,337	690	460	169
Number of IRFs frequently used	1,020	982	556	407	197
Mean number of "frequently used" IRFs by each hospital (SD):	1.16 (0.50)	1.17 (0.51)	1.11 (0.41)	1.16 (0.49)	1.35 (0.70)

Supplemental Table 3.3 (cont'd)

SNF referral networks					
Number of hospitals with at least 1 frequently used SNF	787	752	354	267	117
Number of SNFs frequently used	1,196	1,163	511	424	230
Mean number of frequently used SNFs by each hospital (SD):	1.57 (0.99)	1.56 (1.01)	1.46 (0.84)	1.60 (0.92)	1.97 (1.16)

*Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Typical hospitals: statistically insignificant random intercepts based on the hierarchical logistic regression model *% IRF discharge: Proportion of patients discharged to an IRF versus SNF. Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator

Supplemental Table 3.4: Number of hospitals with regular use and frequently used IRF and SNF referral networks

	All Hospitals		**"Typical Hospitals"**		
Acute care hospitals (n= hospitals)	All (n=3,039)	≥20 cases (n=1,816)	≥20 cases (n=891)	≥50 cases (n=475)	≥100 cases (n=169)
Number of patients (i.e. cases)	145,984	135,415	60,529	47,326	25,980
Hospitals with at least 1 regular use (i.e. single hospital discharges ≥5 patients to a specific rehabilitation facility) IRF *and* SNF referral network					
Number of hospitals	1,225	1,187	669	441	166
Number of patients (i.e. cases)	108,787	108,150	52,900	44,950	25,582
Hospitals with at least 1 frequent use (i.e. single hospital discharges ≥10 patients to a specific rehabilitation facility) IRF *and* SNF referral network					
Number of hospitals	489	489	280	257	117
Number of patients (i.e. cases)	56,960	56,960	29,832	28,890	18,569

Case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF), Typical hospitals had statistically insignificant random intercepts based on the hierarchical logistic regression model

Supplemental Table 4.1: Baseline characteristics of eligible study populations for three emulated trials that compared stroke rehabilitation at IRFs and SNFs

	trial 1 (n=44,950)	trial 2 (n=34,444)	trial 3 (n=19,161)	p-value
Age	81.5 (8.0)	81.1 (8.0)	80.9 (7.9)	<0.001
Race				<0.001
White	82.4%	82.5%	82.4%	
Black	11.2%	10.9%	10.4%	
Hispanic	3.4%	3.5%	3.5%	
Other	3.0%	3.2%	3.6%	
Female sex	60.7%	59.4%	58.6%	<0.001
Median annual household income (per $1,000) *				<0.001
$<25k	3.7%	3.5%	3.1%	
$25-50k	36.7%	34.9%	33.4%	
$50-75k	37.0%	37.6%	37.7%	
$75-100k	13.5%	14.2%	14.8%	
$>100 k	7.3%	7.9%	9.2%	
Missing	1.8%	1.9%	1.9%	
Prior health care utilization*				
Pre-stroke home-time	358.5 (21.4)	359.5 (18.3)	360.1 (16.9)	<0.001
Number of hospitalizations	0.3 (0.7)	0.3 (0.7)	0.3 (0.7)	<0.001
SNF use	11.5%	9.9%	9.0%	<0.001
IRF use	2.3%	2.4%	2.3%	0.68
Comorbidities:				
Elixhauser comorbidity index	4.0 (1.8)	4.0 (1.8)	4.0 (1.8)	0.97
Dementia	9.1%	8.0%	7.4%	<0.001
Stroke Characteristics				
Stroke subtype				0.80
Ischemic	90.9%	90.9%	90.8%	
Intracerebral hemorrhagic	9.1%	9.1%	9.2%	
Stroke administrative severity index				0.99
Mild	38.8%	38.7%	38.8%	
Moderate	39.4%	39.5%	39.3%	
Severe	21.7%	21.8%	21.8%	

Supplemental Table 4.1 (cont'd)

Hospital health services use				
Length of stay (days)	5.2 (2.7)	5.2 (2.7)	5.2 (2.7)	0.98
ICU use	57.2%	56.6%	55.3%	<0.001
Emergency department admission	89.2%	90.9%	91.9%	<0.001
Lifesaving procedures				
Hemodialysis	1.3%	1.0%	1.0%	<0.001
GI tube	6.4%	5.4%	5.1%	<0.001
CPR	0.1%	0.1%	0.0%	0.28
Nutrition	3.6%	3.3%	3.0%	<0.001
Intubation/ventilation	1.9%	1.7%	1.6%	0.026
tPA	7.1%	7.5%	7.8%	0.006
Number of PT CPT revenue codes				<0.001
0	2.2%	1.9%	1.6%	
1-3	36.9%	36.9%	37.8%	
4-7	37.1%	37.6%	37.2%	
8-11	14.5%	14.4%	14.2%	
>11	9.2%	9.2%	9.2%	
Number of OT CPT revenue codes				<0.001
0	16.9%	16.1%	16.4%	
1-2	29.9%	30.3%	31.7%	
3-6	36.4%	36.8%	35.9%	
7-9	9.8%	9.8%	9.5%	
>9	7.0%	6.9%	6.5%	
Number of SLT CPT revenue codes				0.31
0	22.5%	22.6%	22.8%	
1-2	34.3%	34.8%	35.1%	
3-5	29.8%	29.7%	29.2%	
6-7	7.1%	6.8%	6.7%	
>7	6.3%	6.1%	6.3%	
Hospital charge data				<0.001
Pharmacy	25.8%	27.1%	26.5%	
Quartile 1	25.3%	25.8%	26.0%	
Quartile 2	24.8%	23.9%	24.3%	
Quartile 3	24.1%	23.2%	23.2%	
Quartile 4				<0.001

Supplemental Table 4.1 (cont'd)

Laboratory	25.3%	26.7%	26.7%	
Quartile 1	25.5%	26.0%	26.8%	
Quartile 2	26.1%	26.0%	26.1%	
Quartile 3	23.1%	21.4%	20.4%	
Quartile 4				
Hospital Services use (yes/no)				
Inhalation therapy	37.4%	36.6%	36.4%	0.017
MRI	69.2%	70.5%	71.0%	<0.001
Operating room	13.6%	12.6%	12.0%	<0.001

*Median annual household income: taken from race matched zip code data Prior health care utilization
* Taken 1 year prior to the indexed stroke event
* Absolute standardized differences >0.1 considered significant
Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator

Supplemental Table 4.2: Mean length of stay (LOS) at first rehabiliation setting among acute stroke patients discharged to an IRF or SNF

	IRF patients		SNF patients	
Discharge setting:	Mean (SD)	Range [IQR]	Mean (SD)	Range [IQR]
Home	14.6 (7.0)	1-120 [9-18]	32.3 (24.1)	1-352 [16-42]
Acute rehospitalization	8.9 (6.9)	1-42 [3-13]	25.4 (35.6)	1-365 [6-32]
SNF	18.9 (7.0)	1-137 [14-23]	32.0 (31.3)	1-365 [10-44]
Remains patient*	0	0	65.8 (60.0	1-365 [28-93]
In facility death	9.0 (7.0)	1-26 [4-13]	22.7 (39.0)	1-265 [4-24]
Other	14.2 (8.4)	0-54 [8-20]	38.3 (38.3)	1-266 [16-53]

Abbreviations: IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility
Remain patient: Became long term nursing care patient (indicated by “still a patient” discharge code)
Mean LOS calculated by the differences between he admission and discharge date

Supplemental Table 4.3: Number of patients, hospitals, and rehabilitation facilities available for each emulated trial to compare stroke rehabilitation at IRFs and SNFs

	Trial 1		Trial 2		Trial 3	
	All available patients	Matched patients	All available patients	Matched patients	All available patients	Matched patients
Total number of patients	44,950		34,444		19,161	
Number of IRF patients	21,301	11,784	20,588	7,578	12,658	3,728
Number of SNF patients	23,649	11,784	13,856	7,578	6,504	3,728
Number of Hospitals	441	441	441	441	257	257
Number of IRFs	745	662	460	443	297	254
Number of SNFs	5,974	4,579	1,338	1,319	415	414

Matched: Patients matched based on their propensity score estimated using a logistic regression model that adjusted for patient level covariates
Abbreviations: IRF: Inpatient Rehabilitation Facility, SNF: Skilled Nursing Facility

Supplemental Table 4.4: Differences in baseline patient level characteristics between IRF and SNF patients for each matched sample used in the three emulated trials to compared stroke rehabilitation at IRFs and SNFs

	Trial 1 (N=11,784 pairs)			Trial 2 (n=7,578 pairs)			Trial 3 (n=3,728)		
	IRF patients	SNF patients	ASDs	IRF patient	SNF patients	ASDs	IRF patient	SNF patients	ASDs
Age	81.1 (7.3)	81.3 (7.6)	0.02	81.9 (7.0)	82.1 (7.5)	0.02	82.2 (7.0)	82.5 (7.3)	0.03
Race			0.01			0.04			0.05
White	83.2%	82.8%		83.2%	82.7%		82.7%	82.6%	
Black	10.5%	10.8%		9.9%	10.7%		9.6%	10.4%	
Hispanic	3.3%	3.4%		3.5%	3.6%		3.7%	3.8%	
Other	3.0%	3.0%		3.5%	3.0%		3.9%	3.2%	
Female sex	59.9%	60.8%	0.02	60.1%	62.2%	0.04	60.0%	62.4%	0.05
Median annual household income (per $1,000) *			0.02			0.03			0.03
$<25k	3.6%	3.8%		3.3%	3.4%		2.9%	3.0%	
$25-50k	36.4%	36.9%		34.1%	33.5%		31.7%	31.3%	
$50-75k	37.7%	37.4%		37.6%	38.6%		38.6%	39.6%	
$75-100k	13.3%	13.3%		14.7%	15.0%		15.3%	15.3%	
$>100 k	7.1%	6.9%		8.4%	7.9%		10.0%	9.3%	
Missing	1.8%	1.7%		1.8%	1.7%		1.6%	1.5%	
Prior health care utilization*									
Pre-stroke home-time	361.8 (11.2)	361.5 (12.1)	0.03	361.3 (12.6)	360.8 (13.6)	0.04	361.4 (12.1)	361.0 (12.8)	0.02
Number of hospitalizations	0.2 (0.6)	0.2 (0.6)	0.02	0.3 (0.7)	0.3 (0.7)	0.03	0.3 (0.7)	0.3 (0.6)	0.01
SNF use	6.0%	6.7%	0.02	7.1%	8.0%	0.02	7.6%	8.2%	0.02
IRF use	2.2%	2.3%	0.01	2.1%	2.4%	0.03	2.1%	1.7%	0.03

Supplemental Table 4.4 (cont'd)

Comorbidities:									
Elixhauser comorbidity index	4.0 (1.8)	4.0 (1.9)	<0.01	4.1 (1.9)	4.1 (1.9)	<0.01	4.1 (1.9)	4.1 (1.9)	0.01
Dementia	5.4%	5.4%	<0.01	6.3%	6.8%	0.02	7.0%	7.5%	0.02
Stroke Characteristics									
Stroke subtype			<0.01			0.02			0.03
Ischemic	90.7%	90.7%		91.1%	90.6%		91.6%	90.7%	
Intracerebral hemorrhagic	9.3%	9.3%		8.9%	9.4%		8.4%	9.3%	
Stroke administrative severity index			0.01			0.02			0.04
Mild	38.7%	38.5%		38.7%	38.9%		39.9%	38.5%	
Moderate	39.3%	39.6%		39.1%	39.7%		37.6%	39.4%	
Severe	22.0%	22.0%		22.3%	21.5%		22.5%	22.1%	
Hospital health services use									
Length of stay (days)	5.2 (2.7)	5.2 (2.7)	<0.01	5.2 (2.7)	5.2 (2.7)	0.01	5.2 (2.7)	5.2 (2.7)	0.01
ICU use	57.7%	58.2%	0.01	55.1%	55.9%	0.02	53.9%	53.4%	0.01
Emergency department admission	88.6%	88.7%	<0.01	92.8%	92.4%	0.02	94.9%	94.4%	0.03
Lifesaving procedures									
Hemodialysis	1.1%	1.2%	0.01	1.0%	1.1%	0.01	1.2%	1.2%	<0.01
GI tube	5.2%	5.1%	<0.01	5.9%	5.8%	0.01	5.6%	5.7%	0.01
CPR	0.0%	0.1%	<0.01	0.0%	0.0%	<0.01	0.0%	0.0%	0.02
Nutrition	3.5%	3.4%	0.01	3.5%	3.5%	<0.01	3.1%	3.1%	0.01
Intubation/ventilation	2.0%	2.1%	<0.01	1.5%	1.6%	0.01	1.1%	1.2%	0.01
tPA	6.8%	7.0%	0.01	6.8%	6.5%	0.01	6.7%	6.7%	<0.01

Supplemental Table 4.4 (cont'd)

Number of PT CPT revenue codes			<0.03			0.04			0.06
0	1.3%	1.3%		1.7%	1.7%		1.6%	1.6%	
1-3	35.5%	36.7%		36.2%	36.1%		36.3%	38.7%	
4-7	38.0%	37.4%		36.6%	37.6%		36.9%	36.3%	
8-11	15.3%	15.2%		15.0%	14.9%		14.7%	14.2%	
>11	9.9%	9.4%		10.6%	9.6%		10.5%	9.2%	
Number of OT CPT revenue codes			0.01			0.03			0.02
0	13.4%	13.6%		16.3%	15.6%		17.4%	17.1%	
1-2	30.3%	30.6%		30.3%	31.0%		32.1%	32.9%	
3-6	38.1%	38.1%		35.5%	36.3%		34.8%	34.5%	
7-9	10.3%	10.0%		10.3%	9.9%		9.4%	9.3%	
>9	7.9%	7.7%		7.6%	7.2%		6.3%	6.1%	
Number of SLT CPT revenue codes			0.02			0.02			0.05
0	21.5%	22.1%		23.4%	24.0%		24.0%	25.1%	
1-2	33.9%	33.9%		33.3%	33.5%		33.7%	34.1%	
3-5	30.5%	30.2%		29.1%	29.0%		28.4%	28.4%	
6-7	7.2%	7.1%		7.4%	7.0%		7.2%	6.3%	
>7	6.8%	6.7%		6.8%	6.4%		6.8%	6.2%	
Hospital charge data									
Pharmacy			0.01			0.03			0.03
Quartile 1	26.2%	25.8%		26.1%	25.7%		24.8%	25.2%	
Quartile 2	25.8%	25.8%		25.4%	26.7%		25.6%	26.3%	
Quartile 3	24.8%	25.0%		25.1%	25.1%		25.9%	25.9%	
Quartile 4	23.2%	23.5%		23.3%	22.6%		23.7%	22.6%	
Laboratory			0.01			0.01			0.02
Quartile 1	25.5%	25.0%		24.8%	24.6%		23.8%	23.5%	
Quartile 2	26.2%	26.4%		26.0%	25.9%		25.7%	26.4%	
Quartile 3	26.2%	26.2%		26.6%	27.1%		28.1%	28.4%	
Quartile 4	22.1%	22.4%		22.6%	22.5%		22.3%	21.7%	

Supplemental Table 4.4 (cont'd)

Hospital Services use (yes/no)									
Inhalation therapy	37.2%	37.5%	0.01	36.9%	37.0%	<0.01	36.8%	36.5%	0.01
MRI	70.9%	70.3%	0.01	71.0%	70.3%	0.01	69.3%	70.0%	0.02
Operating room	13.6%	13.6%	<0.01	12.5%	12.5%	<0.01	12.2%	12.5%	0.01

*ASDs: Absolute standardized differences=values >0.1 were considered significant

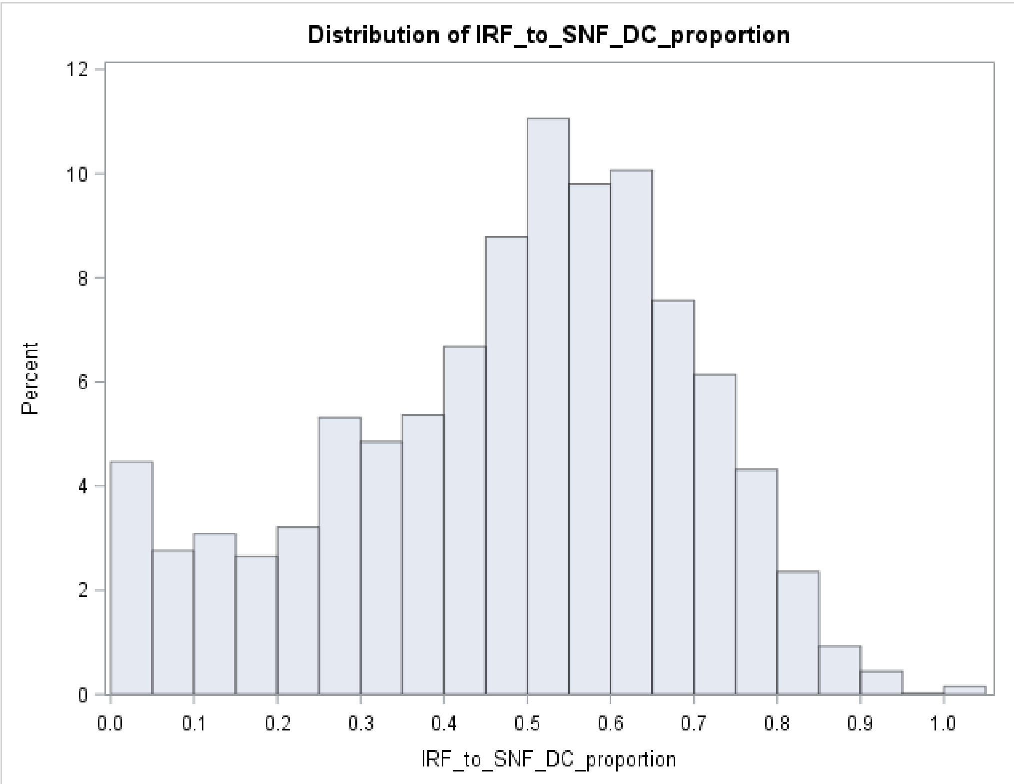
*Median annual household income: taken from race matched zip code data Prior health care utilization

* Taken 1 year prior to the indexed stroke event

* Absolute standardized differences >0.1 considered significant

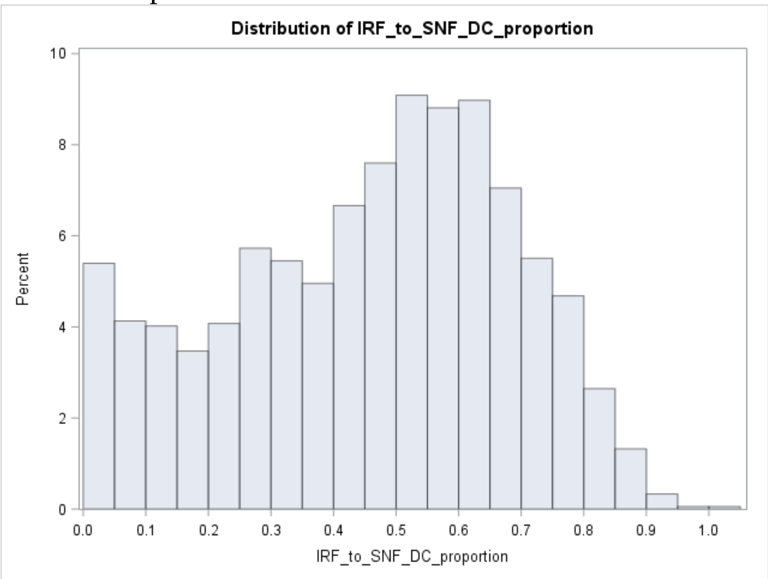
Abbreviations: SNF: Skilled Nursing Facility, IRF: Inpatient Rehabilitation Facility, LOS: Length of Stay, ICU: Intensive Care Unit, tPA: Tissue plasminogen activator

Appendix B: Supplemental Figures

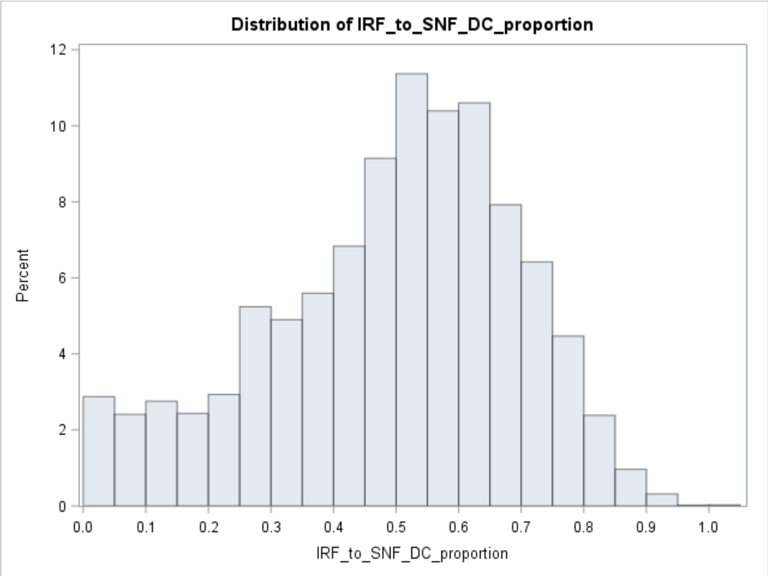


Supplemental Figure 3.1: Patient level variation in the proportion of patients (i.e. cases) discharged to an inpatient rehabilitation facility (IRF) compared to a skilled nursing facility (SNF) among the 1,816 hospitals with at least 20 cases

Panel 1: Hospital-level variation



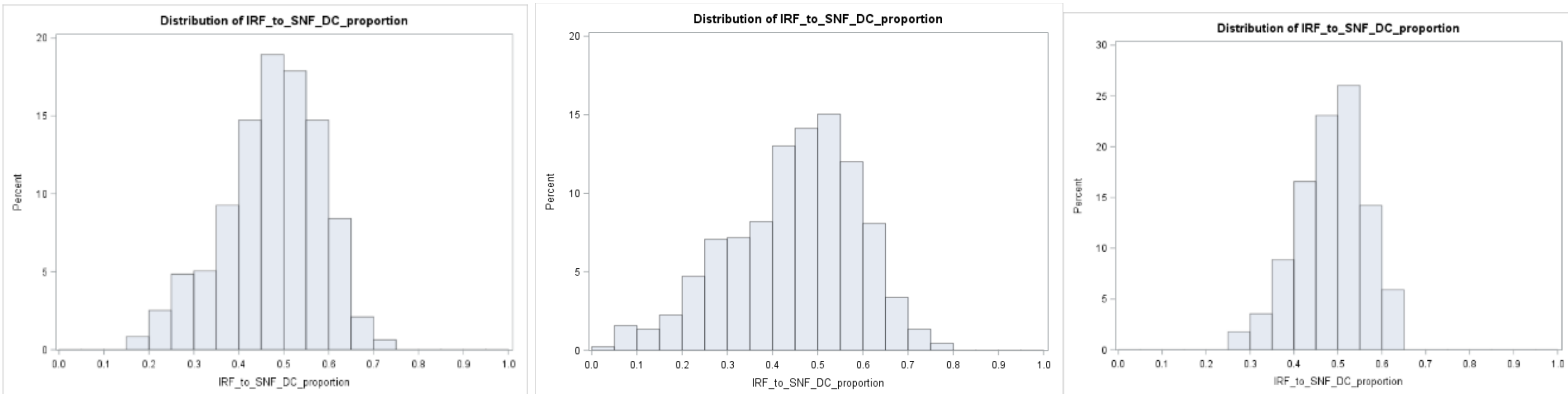
Panel 2: Patient level variation



*case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF)

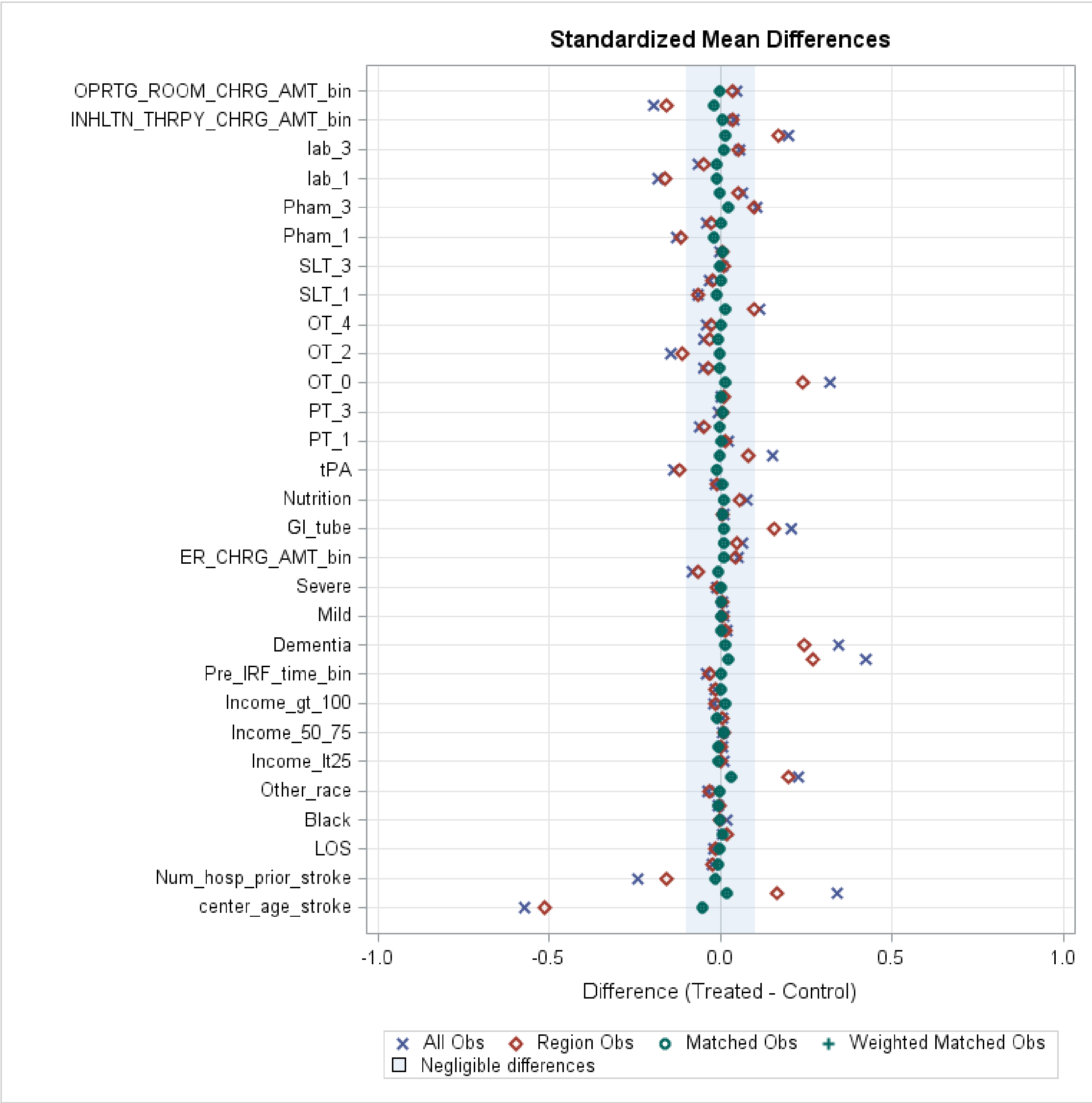
Supplemental Figure 3.2: Hospital and patient level variation in the proportion of patients (i.e. cases) discharged to an inpatient rehabilitation facility (IRF) compared to a skilled nursing facility (SNF) among the 1,816 hospitals with at least 20 cases

Panel 1: Typical hospitals with >20 cases   Panel 2: Typical hospital with >50 cases   Panel 3: Typical hospital with >100 cases



*case: Acute stroke patients discharged to an Inpatient Rehabilitation Facility (IRF) or Skilled Nursing Facility (SNF)
*Typical hospitals had statistically insignificant (p>0.01) random intercepts from on the hierarchical logistic regression model
Panel 1: 891 hospitals and 60,529 patients
Panel 2: 479 hospitals and 47,326 patients
Panel 3: 169 hospitals and 25,980 patients

Supplemental Figure 3.3: Hospital-level variation in the proportion of patients (i.e. cases) discharged to an inpatient rehabilitation facility (IRF) compared to a skilled nursing facility (SNF) reported at the hospital level among patients at typical hospitals



Shaded zone in middle represents zone of clinically irrelevant differences in standardized differences (i.e., >-0.1 and <0.1)

Supplemental Figure 4.1: Standardized differences of patient level covariates for the sensitivity trial after Inpatient Rehabilitation Facility and Skilled Nursing Facility patients were matched across hospitals based on their estimated propensity score

Appendix C: IRB determination

MICHIGAN STATE UNIVERSITY

DETERMINED NOT "HUMAN SUBJECTS"
Revised Common Rule

December 4, 2019

To: Kent Patrick Simmonds

Re: **MSU Study ID:** STUDY00002495
Principal Investigator: Kent Patrick Simmonds
Determination Date: 12/4/2019

Title: Using Administrative Claims Data to Design and Emulate a Trial that Compares Rehabilitation at Inpatient Rehabilitation Facilities Compared to Skilled Nursing Facilities Among Acute Stroke Patients

The activity described in this submission was determined not to involve "human subjects" as defined by the Common Rule as codified in the U.S. Department of Health and Human Services (DHHS) regulations for the protection of human research subjects.

Definition of Human Subject
For DHHS, "*Human subject* means a living individual about whom an investigator (whether professional or student) conducting research:

> (i) Obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or
>
> (ii) Obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens." [45 CFR 46.102(e)(1)]

Determination
This study will conduct a secondary analysis of de-identified Medicare claims data. No additional data will be collected through any interactions or interventions with human subjects.

Hence, the activity does not involve human subjects.

Therefore, the federal regulations for the protection of human subjects would not apply to this activity and Michigan State University (MSU) Institutional Review Board (IRB) approval is not needed to proceed. However, please note that while MSU IRB approval is not required, other federal, state, or local regulations or requirements or ethical or professional standards may still be applicable based on the activity.

Modifications: If any of the activities described in this submission change, please contact the IRB office as the activity may involve human subject research and require IRB approval. For example, this determination is not applicable to activities that may

Office of Regulatory Affairs
Human Research Protection Program

4000 Collins Road
Suite 136
Lansing, MI 48910

517-355-2180
Fax: 517-432-4503
Email: irb@msu.edu
www.hrpp.msu.edu

be regulated by U.S. Food & Drug Administration (FDA), such as those involving drugs, medical devices, human food additives, color additives, electronic products, or any other test articles regulated by the FDA.

Modifications to Funding: **Changes in funding may alter this determination.** For example, MSU IRB review and approval is required if MSU receives an award through a grant, contract, or cooperative agreement directly from a federal agency, even where all non-exempt research involving human subjects are carried out by employees or agents of another institution. In addition, the new funding source may have additional or different requirements.

For More Information: See HRPP Manual Section 4-3, Determination of Human Subject Research (available at hrpp.msu.edu).

Contact Information: If we can be of further assistance or if you have questions, please contact us at 517-355-2180 or via email at IRB@msu.edu. Please visit hrpp.msu.edu to access the HRPP Manual, templates, etc.

REFERENCES

REFERENCES

1. Benjamin EJ, Muntner P, Alonso A, et al. *Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association*. Vol 139.; 2019. doi:10.1161/CIR.0000000000000659

2. Smith EE, Shobha N, Dai D, et al. Risk score for in-hospital ischemic stroke mortality derived and validated within the get with the guidelines-stroke program. *Circulation*. 2010;122(15):1496-1504. doi:10.1161/CIRCULATIONAHA.109.932822

3. Feigin VL, Lawes CMM, Bennett DA, Anderson CS. Stroke epidemiology: A review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol*. 2003;2(1):43-53. doi:10.1016/S1474-4422(03)00266-7

4. Winstein CJ, Stein J, Arena R, et al. *Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association*. Vol 47.; 2016. doi:10.1161/STR.0000000000000098

5. Prvu Bettger J, McCoy L, Smith EE, Fonarow GC, Schwamm LH, Peterson ED. Contemporary trends and predictors of postacute service use and routine discharge home after stroke. *J Am Heart Assoc*. 2015;4(2):1-11. doi:10.1161/JAHA.114.001038

6. Skolarus LE, Feng C, Burke JF. No Racial Difference in Rehabilitation Therapy Across All Post-Acute Care Settings in the Year Following a Stroke. 2017. doi:10.1161/STROKEAHA.117.017290

7. DeJong G, Hsieh CH, Gassaway J, et al. Characterizing Rehabilitation Services for Patients With Knee and Hip Replacement in Skilled Nursing Facilities and Inpatient Rehabilitation Facilities. *Arch Phys Med Rehabil*. 2009;90(8):1269-1283. doi:10.1016/j.apmr.2008.11.021

8. Kane RL. Assessing the Effectiveness of Postacute Care Rehabilitation. *Arch Phys Med Rehabil*. 2007;88(11):1500-1504. doi:10.1016/j.apmr.2007.06.015

9. Reistetter TA, Kuo Y-F, Karmarkar AM, et al. Geographic and facility variation in inpatient stroke rehabilitation: multilevel analysis of functional status. *Arch Phys Med Rehabil*. 2015;96(7):1248-1254. doi:10.1016/j.apmr.2015.02.020

10. Deutsch A, Granger C V., Heinemann AW, et al. Poststroke rehabilitation: Outcomes and reimbursement of inpatient rehabilitation facilities and subacute rehabilitation programs. *Stroke*. 2006;37(6):1477-1482. doi:10.1161/01.STR.0000221172.99375.5a

11. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*.

2011;377(9778):1693-1702. doi:10.1016/S0140-6736(11)60325-5

12. Nguyen VQC, PrvuBettger J, Guerrier T, et al. Factors Associated With Discharge to Home Versus Discharge to Institutional Care After Inpatient Stroke Rehabilitation. *Arch Phys Med Rehabil*. 2015;96(7):1297-1303. doi:10.1016/j.apmr.2015.03.007

13. Medicare Payment Advisory Comission. Report to Congress, March 2017. http://medpac.gov/docs/default-source/reports/mar17_entirereport.pdf. Accessed September 1, 2019.

14. Ackerly D, Grabowski D. Post-Acute Care Reform — Beyond the ACA. *New Engl J Medicice*. 2014;370(8):689-691.

15. Xian Y, Thomas L, Liang L, et al. Unexplained Variation for Hospitals' Use of Inpatient Rehabilitation and Skilled Nursing Facilities After an Acute Ischemic Stroke. *Stroke*. 2017;48(10):2836-2842. doi:10.1161/STROKEAHA.117.016904

16. Freburger J, Holmes G, Ku L-J, Cutchin M, Hearwole-Shank K, Edwards L. Disparities in Post-Acute Rehabilitation Care for Stroke. 2011;92(8):1220-1229. doi:10.1007/s11103-011-9767-z.Plastid

17. Hong I, Karmarkar A, Chan W, et al. Discharge Patterns for Ischemic and Hemorrhagic Stroke Patients Going from Acute Care Hospitals to Inpatient and Skilled Nursing Rehabilitation. *Am J Phys Med Rehabil*. 2018;97(9):636-645. doi:10.1097/PHM.0000000000000932

18. Kramer A, Holthaus D, Goodrish G, Epstein A. *A Study of Stroke Post - Acute Care Costs and Outcomes*.; 2006.

19. Jette DU, Grover L, Keck CP. A qualitative study of clinical decision making in recommending discharge placement from the acute care setting. *Phys Ther*. 2003;83(3):224-236. doi:10.1093/ptj/83.3.224

20. Heinemann AW. State of the science on postacute rehabilitation: Setting a Research Agenda and developing an evidence base for practice and public policy. An introduction. *J Neuroeng Rehabil*. 2007;43:43-48. doi:10.1080/10400435.2008.10131932

21. MedPAC. *Medicare Payment Policy*. Vol March.; 2016.

22. Buntin MB, Garten AD, Paddock S, Saliba D, Totten M, Escarce JJ. How much is postacute care use affected by its availability? *Health Serv Res*. 2005;40(2):413-434. doi:10.1111/j.1475-6773.2005.0i366.x

23. Reistetter TA, Karmarkar AM, Graham JE, et al. Regional variation in stroke rehabilitation outcomes. *Arch Phys Med Rehabil*. 2014;95(1):29-38. doi:10.1016/j.apmr.2013.07.018

24. Burke RE, Juarez-Colunga E, Levy C, Prochazka A V., Coleman EA, Ginde AA. Patient and hospitalization characteristics associated with increased postacute care facility discharges from US hospitals. *Med Care*. 2015;53(6):492-500. doi:10.1097/MLR.0000000000000359

25. Science NA of. *Variation in Healthcare Spending Target Decision Making Not Geography*.; 2010.

26. Newhouse J, Garber A, Graham R, McCoy M, Mancher M, Kibria A. *Variation in Health Care Spending: Target Decision Making, Not Geography*.; 2013. doi:10.17226/18393

27. Alcusky M, Ulbricht CM, Lapane KL. Postacute Care Setting , Facility Characteristics , and Poststroke Outcomes : A Systematic Review. *Arch Phys Med Rehabil*. 2017. doi:10.1016/j.apmr.2017.09.005

28. Deutsch A. Does postacute care site matter? A longitudinal study assessing functional recovery after a stroke. *Arch Phys Med Rehabil*. 2013;94(4):630-632. doi:10.1016/j.apmr.2012.09.033

29. McDougall J, Wright V, Rosenbaum P. The ICF model of functioning and disability: Incorporating quality of life and human development. *Dev Neurorehabil*. 2010;13(3):204-211. doi:10.3109/17518421003620525

30. Chan L. The State-of-the-Science: Challenges in Designing Postacute Care Payment Policy. *Arch Phys Med Rehabil*. 2007;88(11):1522-1525. doi:10.1016/j.apmr.2007.05.032

31. INSPECTOR GENERAL REHABILITATION FACILITY STAYS DID NOT M EET MEDICARE COVERAGE AND DOCUMENTATION Office of Inspector General. 2018;(September).

32. Lee AJ, Huber JH, Stason WB. Factors contributing to practice variation in post-stroke rehabilitation. *Health Serv Res*. 1997;32(2):197-221; discussion 223-227. doi:10.2166/wst.2014.204

33. Chen CC, Heinemann AW, Granger C V., Linn RT. Functional gains and therapy intensity during subacute rehabilitation: A study of 20 facilities. *Arch Phys Med Rehabil*. 2002;83(11):1514-1523. doi:10.1053/apmr.2002.35107

34. Jette DU, Warren RL, Wirtalla C. The relation between therapy intensity and outcomes of rehabilitation in skilled nursing facilities. *Arch Phys Med Rehabil*. 2005;86(3):373-379. doi:10.1016/j.apmr.2004.10.018

35. Bhogal SK, Teasell R, Speechley M. Intensity of aphasia therapy, impact on recovery. *Stroke*. 2003;34(4):987-992. doi:10.1161/01.STR.0000062343.64383.D0

36. Lee KB, Lim SH, Kim KH, et al. Six-month functional recovery of stroke patients: A multi-time-point study. *Int J Rehabil Res*. 2015;38(2):173-180. doi:10.1097/MRR.0000000000000108

37. Chan L, Sandel ME, Jette AM, et al. Does postacute care site matter? A longitudinal study assessing functional recovery after a stroke. *Arch Phys Med Rehabil*. 2013;94(4):622-629. doi:10.1016/j.apmr.2012.09.033

38. Anderson C, Mhurchu CN, Rubenach S, Clark M, Spencer C, Winsor A. Home or hospital for stroke rehabilitation? Results of a randomized controlled trial. II: Cost minimization analysis at 6 months. *Stroke*. 2000;31(5):1032-1037. doi:10.1161/01.STR.31.5.1032

39. Rønning OM, Guldvog B. Stroke units versus general medical wards, I: Twelve-and eighteen-month survival: A randomized, controlled trial. *Stroke*. 1998;29(1):58-62. doi:10.1161/01.STR.29.1.58

40. Berges IM, Kuo YF, Ottenbacher KJ, Seale GS, Ostir G V. Recovery of Functional Status After Stroke in a Tri-Ethnic Population. *PM R*. 2012;4(4):290-295. doi:10.1016/j.pmrj.2012.01.010

41. Stein J, Bettger JP, Sicklick A, Hedeman R, Magdon-Ismail Z, Schwamm LH. Use of a standardized assessment to predict rehabilitation care after acute stroke. *Arch Phys Med Rehabil*. 2015;96(2):210-217. doi:10.1016/j.apmr.2014.07.403

42. Porter M. What Is Value in Health Care? *New Engl J Medicice*. 2010;363:2477-2481.

43. Butzer JF, Kozlowski AJ, Virva R. Measuring Value in Postacute Care. *Arch Phys Med Rehabil*. 2019;100(5):990-994. doi:10.1016/j.apmr.2018.11.008

44. Chen Q, Kane RL, Finch MD. The cost effectiveness of post-acute care for elderly Medicare beneficiaries. *Inquiry*. 2000;37(4):359-375. http://search.proquest.com.ezproxy.gvsu.edu/docview/220999318?accountid=39473%5Cnhttp://vq9xh3gm7u.search.serialssolutions.com/?ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/ProQ:medicalshell&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft.g

45. DaVanzo J, El-Gamil A, Li J, Shimer M, Manolov N, Dobson A. Assessment of Patient Outcomes of Rehabilitative Care Provided in Inpatient Rehabilitation Facilities (IRFs) and After Discharge. 2014:1-62. https://www.amrpa.org/newsroom/Dobson DaVanzo AMRPA 2-page summary REVISED 3.10.14 DATED 7.10.14.pdf.

46. Cormier DJ, Frantz MA, Rand E, Stein J. Physiatrist referral preferences for postacute stroke rehabilitation. *Medicine (Baltimore)*. 2016;95(33):e4356. doi:10.1097/MD.0000000000004356

47. Kane RL, Lin W-C, Blewett LA. Geographic variation in the use of post-acute care.

Health Serv Res. 2002;37(3):667-682. doi:10.1111/1475-6773.00043

48. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464-475. doi:10.1016/j.jclinepi.2008.12.011

49. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

50. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur J Epidemiol*. 2017;32(6):473-475. doi:10.1007/s10654-017-0293-4

51. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*. 2011;46:399-424. doi:10.1080/00273171.2011.568786

52. Garcia-Albeniz X, Hsu J, Hern MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol*. 2017;32(6):495-500. doi:10.1021/acssynbio.5b00266.Quantitative

53. Cain LE, Saag MS, Petersen M, et al. Using observational data to emulate a randomized trial of dynamic treatment- switching strategies : an application to antiretroviral therapy. 2016;(December 2015):2038-2049. doi:10.1093/ije/dyv295

54. Hernan M, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart. *Epidemiology*. 2008;19(6):766-779. doi:10.1097/EDE.0b013e3181875e61.Observational

55. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med*. 2016;375:454-463. doi:10.1056/NEJMra1510059

56. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: Designing trials that are fit for purpose. *BMJ*. 2015;350. doi:10.1136/bmj.h2147

57. Lipman PD, Dluzak L, Stoney CM. Is this study feasible? Facilitating management of pragmatic trial planning milestones under a phased award funding mechanism. *Trials*. 2019;20(1):1-9. doi:10.1186/s13063-019-3387-3

58. Didelez V. Commentary: Should the analysis of observational data always be preceded by specifying a target experimental trial? *Int J Epidemiol*. 2016;45(6):2049-2051. doi:10.1093/ije/dyw032

59. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70-75. doi:10.1016/j.jclinepi.2016.04.014

60. Chaix B, Rosvall M, Lynch J, Merlo J. Disentangling contextual effects on cause-specific

mortality in a longitudinal 23-year follow-up study: Impact of population density or socioeconomic environment? *Int J Epidemiol*. 2006;35(3):633-643. doi:10.1093/ije/dyl009

61. Coles E, Wells M, Maxwell M, et al. The influence of contextual factors on healthcare quality improvement initiatives: What works, for whom and in what setting? Protocol for a realist review. *Syst Rev*. 2017;6(1):1-10. doi:10.1186/s13643-017-0566-8

62. Merlo J, Wagner P, Ghith N, Leckie G. An original stepwise multilevel logistic regression analysis of discriminatory accuracy: The case of neighbourhoods and health. *PLoS One*. 2016;11(4):1-31. doi:10.1371/journal.pone.0153778

63. Merlo J. Invited commentary: Multilevel analysis of individual heterogeneity-a fundamental critique of the current probabilistic risk factor epidemiology. *Am J Epidemiol*. 2014;180(2):208-212. doi:10.1093/aje/kwu108

64. Kent DM, Steyerberg E, Van Klaveren D. Personalized evidence based medicine: Predictive approaches to heterogeneous treatment effects. *BMJ*. 2018;363. doi:10.1136/bmj.k4245

65. Kent DM, Rothwell PM, Ioannidis JPA, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: A proposal. *Trials*. 2010;11:1-10. doi:10.1186/1745-6215-11-85

66. CMS. Medicare standard analytic files. https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/StandardAnalyticalFiles.

67. Census U. American Community Survey. https://www.census.gov/programs-surveys/acs.

68. Centers for Medicare & Medicaid Services. Provider of Service File. https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.

69. Centers for Medicare & Medicaid Services. Hospital Compare Data.

70. Fonarow GC, Liang L, Thomas L, et al. Assessment of Home-Time after Acute Ischemic Stroke in Medicare Beneficiaries. *Stroke*. 2016;47(3):836-842. doi:10.1161/STROKEAHA.115.011599

71. Elixhauser A, Steiner C, Harris DR. Comorbidity Measures for Use with Administrative Data Author (s): Anne Elixhauser , Claudia Steiner , D . Robert Harris and Rosanna M . Coffey Published by : Lippincott Williams & Wilkins Stable URL : http://www.jstor.org/stable/3766985 Accessed : 26-07. *Med Care*. 1998;36(1):8-27.

72. Simpson AN, Wilmskoetter J, Hong I, et al. Stroke Administrative Severity Index : using administrative data for 30-day poststroke outcomes prediction. *J Comp Eff Res*. 2018;7:293-304.

73. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;(July):3083-3107. doi:10.1002/sim

74. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *Eur Urol*. 2015;67(6):1142-1151. doi:10.1016/j.eururo.2014.11.025

75. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 2009;338(7707):1373-1377. doi:10.1136/bmj.b604

76. Cho JS, Hu Z, Fell N, Heath GW, Qayyum R, Sartipi M. Hospital Discharge Disposition of Stroke Patients in Tennessee. *South Med J*. 2017;110(9):594-600. doi:10.14423/SMJ.0000000000000694

77. Hosmer. *Applied Logistic Regression*. 2nd ed. New York, NY: Wiley & Sons, Inc; 2000.

78. Merlo J, Wagner P, Austin PC, Subramanian S V., Leckie G. General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial. *SSM - Popul Heal*. 2018;5(March):33-37. doi:10.1016/j.ssmph.2018.05.006

79. Austin PC, Merlo J. Intermediate and advanced topics in multilevel logistic regression analysis. *Stat Med*. 2017;36(20):3257-3277. doi:10.1002/sim.7336

80. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health*. 2006;60(4):290-297. doi:10.1136/jech.2004.029454

81. Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. *Int J Epidemiol*. 2016;(July):dyw118. doi:10.1093/ije/dyw118

82. Leyland AH, Boddy FA. League tables and acute myocardial infarction. *Lancet*. 1998;351(9102):555-558. doi:10.1016/S0140-6736(97)09362-8

83. Austin PC, Alter DA, Tu J V. The Use of Fixed- and Random-Effects Models for Classifying Hospitals as Mortality Outliers: A Monte Carlo Assessment. *Med Decis Mak*. 2003;23(6):526-539. doi:10.1177/0272989X03258443

84. Austin PC. A comparison of Bayesian methods for profiling hospital performance. *Med Decis Mak*. 2002;22(2):163-172. doi:10.1177/0272989X0202200213

85. Leckie G, Goldstein H. The limitations of using school league tables to inform school

choice. *J R Stat Soc Ser A Stat Soc*. 2009;172(4):835-851. doi:10.1111/j.1467-985X.2009.00597.x

86. Bouwmeester W, Twisk JW, Kappen TH, Van Klei WA, Moons KG, Vergouwe Y. Prediction models for clustered data: Comparison of a random intercept and standard regression model. *BMC Med Res Methodol*. 2013;13(1). doi:10.1186/1471-2288-13-19

87. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med*. 2007;35(9):2052-2056. doi:10.1097/01.CCM.0000275267.64078.B0

88. Tirschwell DL, Longstreth WT. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470. doi:10.1161/01.STR.0000032240.28636.BD

89. Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand SLT. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. *Circulation*. 2007;115(12):1518-1527. doi:10.1161/CIRCULATIONAHA.106.633008

90. Worsley SD, Oude Rengerink K, Irving E, et al. Series: Pragmatic trials and real world evidence: Paper 2. Setting, sites, and investigator selection. *J Clin Epidemiol*. 2017;88:14-20. doi:10.1016/j.jclinepi.2017.05.003

91. Gheorghe A, Roberts TE, Ives JC, Fletcher BR, Calvert M. Centre Selection for Clinical Trials and the Generalisability of Results: A Mixed Methods Study. *PLoS One*. 2013;8(2):1-9. doi:10.1371/journal.pone.0056560

92. Schwartz D, Leelouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis*. 1967;20:637-648.

93. Dennis M, Mead G, Forbes J, et al. Effects of fluoxetine on functional outcomes after acute stroke (FOCUS): a pragmatic, double-blind, randomised, controlled trial. *Lancet*. 2019;393(10168):265-274. doi:10.1016/S0140-6736(18)32823-X

94. Duncan PW, Bushnell CD, Rosamond WD, et al. The Comprehensive Post-Acute Stroke Services (COMPASS) study: Design and methods for a cluster-randomized pragmatic trial. *BMC Neurol*. 2017;17(1):1-13. doi:10.1186/s12883-017-0907-1

95. Ciccone A, Valvassori L, Nichelatti M, et al. Endovascular Treatment for Acute Ischemic Stroke for the SYNTHESIS Expansion Investigators *. 2013;368(10):904-913. doi:10.1056/NEJMoa1213701

96. Mozaffarian D, Benjamin EJ, Go AS, et al. *Heart Disease and Stroke Statistics-2016 Update a Report from the American Heart Association*. Vol 133.; 2016. doi:10.1161/CIR.0000000000000350

97. Bettger JP, Thomas L, Li L. *Comparing Recovery Options for Stroke Patientse*. Washington, DC; 2019. https://doi.org/10.25302/3.2019.CER.130.

98. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions : the new Medical Research Council guidance. *Br Med J*. 2008;1655(September):1-6. doi:10.1136/bmj.a1655

99. Braslow JT, Duan N, Starks SL, Polo A, Bromley E, Wells KB. Generalizability of studies on mental health treatment and outcomes, 1981 to 1996. *Psychiatr Serv*. 2005;56(10):1261-1268. doi:10.1176/appi.ps.56.10.1261

100. Lau C, Alpert A, Huckfeldt P, et al. Post-acute referral patterns for hospitals and implications for bundled payment initiatives. *Healthcare*. 2014;2(3):190-195. doi:10.1016/j.hjdsi.2014.05.004

101. Johnson AM, Jones SB, Duncan PW, et al. Hospital recruitment for a pragmatic cluster-randomized clinical trial: Lessons learned from the COMPASS study. *Trials*. 2018;19(1):1-9. doi:10.1186/s13063-017-2434-1

102. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med*. 2002;137(6):511-520. doi:10.7326/0003-4819-137-6-200209170-00012

103. Nguyen NT, Paya M, Stevens CM, et al. The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. *Ann Surg*. 2004;240(4):586-594. doi:10.1097/01.sla.0000140752.74893.24

104. Gruen R, Pitt V, Green S, Parkhill A, Campbell D, Jolley D. The Effect of Provider Case Volume on Cancer Mortality. *CA Cancer J Clin*. 2009;59(3):192-211. doi:10.3322/caac.20018.Available

105. Fonarow GC, Smith EE, Reeves MJ, et al. Hospital-level variation in mortality and rehospitalization for medicare beneficiaries with acute ischemic stroke. *Stroke*. 2011;42(1):159-166. doi:10.1161/STROKEAHA.110.601831

106. O'Brien EC, Xian Y, Xu H, et al. Hospital Variation in Home-Time after Acute Ischemic Stroke: Insights from the PROSPER Study (Patient-Centered Research into Outcomes Stroke Patients Prefer and Effectiveness Research). *Stroke*. 2016;47(10):2627-2633. doi:10.1161/STROKEAHA.116.013563

107. Middleton A, Graham JE, Bettger JP, Haas A, Ottenbacher KJ. Facility and Geographic Variation in Rates of Successful Community Discharge After Inpatient Rehabilitation Among Medicare Fee-for-Service Beneficiaries. *JAMA Netw Open*. 2018;1(7):1-11. doi:10.1001/jamanetworkopen.2018.4332

108. Graham JE, Deutsch A, O'Connell AA, Karmarkar AM, Granger C V., Ottenbacher KJ.

Inpatient Rehabilitation Volume and Functional Outcomes in Stroke, Lower Extremity Fracture, and Lower Extremity Joint Replacement. *Med Care*. 2013;51(5):404-412. doi:10.1038/jid.2014.371

109. Graham JE, Prvu Bettger J, Middleton A, Spratt H, Sharma G, Ottenbacher KJ. Effects of Acute–Postacute Continuity on Community Discharge and 30-Day Rehospitalization Following Inpatient Rehabilitation. *Health Serv Res*. 2017;52(5):1631-1646. doi:10.1111/1475-6773.12678

110. Rahman M, Foster AD, Grabowski DC, Zinn JS, Mor V. Effect of hospital-SNF referral linkages on rehospitalization. *Health Serv Res*. 2013;48(6 PART1):1898-1919. doi:10.1111/1475-6773.12112

111. Whyte J, Hart T. It's More Than a Black Box; It's a Russian Doll. *Am J Phys Med Rehabil*. 2003;82(8):639-652. doi:10.1097/01.phm.0000078200.61840.2d

112. DeJong G, Horn SD, Conroy B, Nichols D, Healton EB. Opening the black box of poststroke rehabilitation: Stroke rehabilitation patients, processes, and outcomes. *Arch Phys Med Rehabil*. 2005;86(12 SUPPL.):1-7. doi:10.1016/S0966-6532(05)80005-8

113. Freedman B. Freedman (1987) Equipoise and the Ethics of Clinical Research.pdf. *New Engl J Medicice*. 1987;317(3):141-145.

114. Meulemeester J De, Fedyk M, Jurkovic L, et al. Many randomized clinical trials may not be justified : a cross-sectional analysis of the ethics and science of randomized clinical trials. *J Clin Epidemiol*. 2018;97:20-25. doi:10.1016/j.jclinepi.2017.12.027

115. Phillips S, John A, Clara L, et al. Is the concept of clinical equipoise still relevant to. 2017;5787(December):10-12. doi:10.1136/bmj.j5787

116. Borschmann R, Patterson S, Poovendran D, Wilson D, Weaver T. Influences on recruitment to randomised controlled trials in mental health settings in England: A national cross-sectional survey of researchers working for the Mental Health Research Network. *BMC Med Res Methodol*. 2014;14(1). doi:10.1186/1471-2288-14-23

117. Patterson S, Mairs H, Borschmann R. Successful recruitment to trials: A phased approach to opening gates and building bridges. *BMC Med Res Methodol*. 2011;11. doi:10.1186/1471-2288-11-73

118. Sood N, Huckfeldt PJ, Escarce JJ, Grabowski DC, Newhouse JP. Medicare's bundled payment pilot for acute and postacute care: Analysis and recommendations on where to begin. *Health Aff*. 2011;30(9):1708-1717. doi:10.1377/hlthaff.2010.0394

119. Admon AJ, Donnelly JP, Casey JD, et al. Emulating a novel clinical trial using existing observational data predicting results of the PREVENT study. *Ann Am Thorac Soc*. 2019;16(8):998-1007. doi:10.1513/AnnalsATS.201903-241OC

120. Wang H, Sandel E, Terdiman J, et al. Postacute Care and Ischemic Stroke Mortality: Findings From an Integrated Health Care System in Northern California. *PM&R*. 2011;3:686-694. doi:10.1016/j.pmrj.2011.06.008

121. Kane RL, Chen Q, Finch M, Blewett L, Burns R, Moskowitz M. The optimal outcomes of post-hospital care under medicare. *Health Serv Res*. 2000;35(3):615-661. http://www.ncbi.nlm.nih.gov/pubmed/10966088%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1089140.

122. Kind AJH, Smith MA, Liou JI, Pandhi N, Frytak JR, Finch MD. Discharge Destination's Effect on Bounce-Back Risk in Black, White, and Hispanic Acute Ischemic Stroke Patients. *Arch Phys Med Rehabil*. 2010;91(2):189-195. doi:10.1016/j.apmr.2009.10.015

123. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat*. 2011;10(2):150-161. doi:10.1002/pst.433

124. Li F, Zaslavsky A, Landrum M. Propensity Score Weighting with Multilevel Data. *Stat Med*. 2013;32(19):3373-3387. doi:10.1002/sim.5786.Propensity

125. Dehejia RH, Wahba S. Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *J Am Stat Assoc*. 1999;94(448):1053-1062. doi:10.1080/01621459.1999.10473858

126. IMPACT Act of 2014 Data Standardization & Cross Setting Measures. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/IMPACT-Act-of-2014/IMPACT-Act-of-2014-Data-Standardization-and-Cross-Setting-Measures. Published 2018.

127. Cary MP, Prvu Bettger J, Jarvis JM, Ottenbacher KJ, Graham JE. Successful Community Discharge Following Postacute Rehabilitation for Medicare Beneficiaries: Analysis of a Patient-Centered Quality Measure. *Health Serv Res*. 2018;53(4):2470-2482. doi:10.1111/1475-6773.12796

128. Cummings P, McKnight B. Analysis of Matched Cohort Data. *Stata J*. 2004;4(3):274-281. doi:10.1177/1536867x0400400305

129. Van Der Weele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

130. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc*. 1958;53(282):457-481. doi:10.1080/01621459.1958.10501452

131. Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in cox regression. *Stat Med*. 1995;14(15):1707-1723.

doi:10.1002/sim.4780141510

132. Austin PC, Small DS. The use of bootstrapping when using propensity-score matching without replacement: A simulation study. *Stat Med*. 2014;33(24):4306-4319. doi:10.1002/sim.6276

133. Xie Y, Brand JE, Jann B. Estimating Heterogeneous Treatment Effects with Observational Data. *Sociol Methodol*. 2012;42(1):314-347. doi:10.1177/0081175012452652

134. Wan F. Matched or unmatched analyses with propensity-score–matched data? *Stat Med*. 2019;38(2):289-300. doi:10.1002/sim.7976

135. Kumar A, Graham JE, Resnik L, et al. Examining the Association Between Comorbidity Indexes and Functional Status in Hospitalized Medicare Fee-for-Service Beneficiaries. *Phys Ther*. 2016;96(2):232-240. doi:10.2522/ptj.20150039

136. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. *J Gen Intern Med*. 2011;26(2):192-196. doi:10.1007/s11606-010-1513-8

137. Brand JE, Xie Y. Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *Am Sociol Rev*. 2010;75(2):273-302. doi:10.1177/0003122410363567

138. PCORI. Guidance on the Design and Conduct of Trials in Real-World Settings: Facotrs to Consider in Pragmatic Patient-Centered Outcomes Research. https://www.pcori.org/sites/default/files/PCORI-Guidance-Design-and-Conduct-of-Trials-Real-World-Settings-Factors-to-Consider-Pragmatic-PCOR.pdf.

139. PCORI Methodolgy Comittee. *The Patient-Centered Outcomes Research Institute (PCORI) Methodology Report*.; 2019.

140. Weisscher N, Vermeulen M, Roos YB, De Haan RJ. What should be defined as good outcome in stroke trials; A modified Rankin score of 0-1 or 0-2? *J Neurol*. 2008;255(6):867-874. doi:10.1007/s00415-008-0796-8

141. Williams BC, Li Y, Fries BE, Warren RL. Predicting patient scores between the functional independence measure and the minimum data set: Development and performance of a FIM-MDS “crosswalk.” *Arch Phys Med Rehabil*. 1997;78(1):48-54. doi:10.1016/S0003-9993(97)90009-5

142. Murray PK, Dawson N V., Thomas CL, Cebul RD. Are we selecting the right patients for stroke rehabilitation in nursing homes? *Arch Phys Med Rehabil*. 2005;86(5):876-880. doi:10.1016/j.apmr.2004.10.045

143. Wang Y, Cai H, Li C, et al. Optimal caliper width for propensity score matching of three treatment groups: A Monte Carlo study. *PLoS One*. 2013;8(12).

doi:10.1371/journal.pone.0081045

144. Ottenbacher KJ, Graham JE. The State-of-the-Science: Access to Postacute Care Rehabilitation Services. A Review. *Arch Phys Med Rehabil*. 2007;88(11):1513-1521. doi:10.1016/j.apmr.2007.06.761

145. Hakkennes SJ, Brock K, Hill KD. Selection for inpatient rehabilitation after acute stroke: A systematic review of the literature. *Arch Phys Med Rehabil*. 2011;92(12):2057-2070. doi:10.1016/j.apmr.2011.07.189

146. Mees M, Klein J, Yperzeele L, Vanacker P, Cras P. Predicting discharge destination after stroke: A systematic review. *Clin Neurol Neurosurg*. 2016;142:15-21. doi:10.1016/j.clineuro.2016.01.004

147. Gifford F. Community-equipoise and the ethics of randomized control trials. *Bioethics*. 1995;9(2):127-148.

148. Parsons NR, Kulikov Y, Girling A, Griffin D. A statistical framework for quantifying clinical equipoise for individual cases during randomized controlled surgical trials. *Trials*. 2011;12:1-11. doi:10.1186/1745-6215-12-258

149. Selker HP, Ruthazer R, Terrin N, Griffith JL, Concannon T, Kent DM. Random treatment assignment using mathematical equipoise for comparative effectiveness trials. *Clin Transl Sci*. 2011;4(1):10-16. doi:10.1111/j.1752-8062.2010.00253.x

150. Kang M, Ragan BG, Park JH. Issues in outcomes research: An overview of randomization techniques for clinical trials. *J Athl Train*. 2008;43(2):215-221. doi:10.4085/1062-6050-43.2.215

151. Jette AM, Haley SM, Tao W, et al. Prospective Evaluation of the AM-PAC in Outpatient Settings. 2007;87(4).

152. Jette DU, Stilphen M, Ranganathan VK, Passek SD, Frost FS, Jette AM. Validity of the AM-PAC "6-Clicks" Inpatient Daily Activity and Basic Mobility Short Forms. *Phys Ther*. 2014;94(3):379-391.

153. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *Br Med J*. 2000;321:694-969.

154. Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA - J Am Med Assoc*. 2011;305(23):2462-2463. doi:10.1001/jama.2011.822

155. Haut ER, Pronovost PJ, Schneider EB. Limitations of administrative databases. *JAMA - J Am Med Assoc*. 2012;307(24):2589. doi:10.1001/jama.2012.6626

156. Hennessy M, Hunter A, Healy P, Galvin S, Houghton C. Improving trial recruitment processes: How qualitative methodologies can be used to address the top 10 research

priorities identified within the PRioRiTy study. *Trials*. 2018;19(1):1-5. doi:10.1186/s13063-018-2964-1

157. Buntin M, Carrie C, Deb P, Sood N, Escarce J. Medicare Spending and Outcomes After Postacute Care for Stroke and Hip Fracture. *Med Care*. 2010;48(9):776-784.