NIH Public Access
Author Manuscript
*Med Care*. Author manuscript; available in PMC 2014 May 01.

Published in final edited form as:
*Med Care*. 2013 May ; 51(5): 404–412. doi:10.1097/MLR.0b013e318286e3c8.

# Inpatient Rehabilitation Volume and Functional Outcomes in Stroke, Lower Extremity Fracture, and Lower Extremity Joint Replacement

**James E. Graham, PhD, DC**, **Anne Deutsch, RN, PhD, CRRN**, **Ann A. O'Connell, EdD**, **Amol M. Karmarkar, PhD, OTR**, **Carl V. Granger, MD**, and **Kenneth J. Ottenbacher, PhD, OTR**
Division of Rehabilitation Sciences, University of Texas Medical Branch, Galveston, TX (Graham, Karmarkar, and Ottenbacher); Rehabilitation Institute of Chicago, Chicago, IL (Deutsch); Northwestern University Feinberg School of Medicine, Department of Physical Medicine and Rehabilitation and Institute for Healthcare Studies (Deutsch); RTI International (Deutsch); College of Education and Human Ecology, The Ohio State University, Columbus, OH (O'Connell); and Uniform Data System for Medical Rehabilitation, University at Buffalo, Buffalo, NY (Granger)

## Abstract

**Background—**It is unclear if volume-outcome relationships exist in inpatient rehabilitation.

**Objectives—**Assess associations between facility volumes and two patient-centered outcomes in the three most common diagnostic groups in inpatient rehabilitation.

**Research Design—**We used hierarchical linear and generalized linear models to analyze administrative assessment data from patients receiving inpatient rehabilitation services for stroke (n=202,423), lower extremity fracture (n=132,194), or lower extremity joint replacement (n=148,068) between 2006 and 2008 in 717 rehabilitation facilities across the U.S. Facilities were assigned to quintiles based on average annual diagnosis-specific patient volumes.

**Measures—**Discharge functional status (FIM instrument) and probability of home discharge.

**Results—**Facility-level factors accounted for 6–15% of the variance in discharge FIM total scores and 3–5% of the variance in home discharge probability across the 3 diagnostic groups. We used the middle volume quintile (Q3) as the reference group for all analyses and detected small, but statistically significant ($p < .01$) associations with discharge functional status in all three diagnosis groups. Only the highest volume quintile (Q5) reached statistical significance, displaying higher functional status ratings than Q3 each time. The largest effect was observed in FIM total scores among fracture patients, with only a 3.6-point difference in Q5 and Q3 group means. Volume was not independently related to home discharge.

**Conclusions—**Outcome-specific volume effects ranged from small (functional status) to none (home discharge) in all three diagnostic groups. Patients with these conditions can be treated locally rather than at higher-volume regional centers. Further regionalization of inpatient rehabilitation services is not needed for these conditions.

Corresponding Author: James E. Graham, University of Texas Medical Branch, Division of Rehabilitation Sciences, 301 University Blvd, Galveston, TX 77555-1137, jegraham@utmb.edu, Phone: 409-747-1634, Fax: 409-747-1638.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors have no conflicts of interest or financial disclosures to report.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Keywords**

rehabilitation; post-acute care; volume

## INTRODUCTION

Since the seminal study by Luft and colleagues[1] showing an association between hospital surgical volume and patient mortality, evidence for the association between volume and outcomes has increased substantially.[2–4] The consistency and scope of the data have led many to recommend that volume is a proxy for quality of care, in the absence of comprehensive provider-specific information.[5–10] Volume is one of three measures included in the Agency for Healthcare Research and Quality's list of "Inpatient Quality Indicators."[11] Missing from this extensive body of literature are data on the volume-outcomes relationship beyond acute care,[12] such as inpatient rehabilitation facilities.

Aspects of the prospective payment system (PPS) for inpatient rehabilitation facilities make rehabilitation volume a particular area of interest. Mallinson et al.[13] reported on trends in facility dynamics over the first three years following introduction of the PPS. Total occupancy rates remained relatively steady while average lengths of stay decreased and total number of beds increased. Maintaining stable occupancy rates despite declining lengths of stay was likely the result of compensatory increases in admissions; i.e. strategies to increase patient volume.[13] However, the "60 percent rule" from the Centers for Medicare and Medicaid Services (CMS) mandates that at least 60% of annual facility discharges must fall within 13 specified diagnoses.[14] Variations in both the ability to increase total patient volume and to comply with the 60 percent rule have led not only to differences in overall volume between facilities, but also to purposeful differences in diagnosis-specific volumes within facilities.

It is currently unknown if inpatient rehabilitation volume is associated with rehabilitation-relevant outcomes. Volume-outcomes research in acute care typically focuses on survival; however, medical rehabilitation is designed to maximize functioning and community independence. Mortality is not an appropriate outcome for rehabilitation research.[15] The purpose of this study was to examine the relationships between diagnosis-specific inpatient rehabilitation volumes and rehabilitation-relevant outcomes: discharge functional status and discharge setting. Our analysis included the three most common inpatient rehabilitation diagnosis categories: stroke, lower extremity fracture, and lower extremity joint replacement.[16]

## METHODS

### Data Source

Data were obtained from 717 rehabilitation hospitals and units across the U.S. that submitted data to the Uniform Data System for Medical Rehabilitation (UDSMR) for 3 consecutive years: 2006–2008. The UDSMR database is the largest nongovernment registry for medical rehabilitation data in the U.S.[17] The database includes the same items that are submitted to CMS as part of the inpatient rehabilitation facility PPS, but the UDSMR database includes data for all patients, regardless of payer. This study was approved by the University's institutional review board.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

## Study Sample

Given the hierarchical nature of our study question, we were interested in modeling both facility- and patient-level effects. Our sample of 717 inpatient rehabilitation facilities represents approximately 60% of all rehabilitation hospitals in the U.S. Patient records for this study were limited to those receiving inpatient rehabilitation services for stroke (impairment group codes 1.1 to 1.9), lower extremity fracture (impairment group codes 8.1 to 8.2), or lower extremity joint replacement (impairment group codes 8.5 to 8.7) and who were discharged between January 1, 2006 and December 31, 2008. The initial sample contained 551,773 patient records. Patients were excluded if it was not their initial rehabilitation admission (n=23,775), they were not between the ages of 45 and 100 years (n=17,489), if they stayed in the rehabilitation facility longer than 60 days (n=8,994), there was more than 60 days between impairment onset and rehabilitation admission (n=12,728), their records were missing critical data (n=5,562), or they died during their stay ((n = 540). The final sample included 482,694 patients, which represents 87% of the original cohort: stroke = 202,423; lower extremity fracture = 132,194; and lower extremity joint replacement = 148,068.

## Independent Variable

**Diagnosis-specific facility volumes—**The total numbers of patients within each of the three diagnosis categories over the three-year study period were determined for all facilities. These summed values were divided by three to compute average annual diagnosis-specific facility volumes. Volume quintiles were established to present descriptive summaries and for inclusion in the predictive models in order to examine the potential for non-linear patterns (e.g. volume thresholds) in the volume-outcomes relationships. A few facilities did not admit patients with lower extremity fracture (n = 4) or joint replacement (n = 5) during the study period. Thus, the facility counts were 713 and 712 for fracture and joint replacement, respectively.

## Dependent Variables

**Discharge Functional Status—**Functional status was evaluated with the 18-item FIM (Functional Independence Measure) instrument, which is part of the inpatient rehabilitation facility–patient assessment instrument (IRF-PAI).[18] The reliability and validity of FIM data have been substantiated.[19;20] The FIM instrument assesses patient abilities across six subscales relating to motor (self care, sphincter control, transfers, and mobility) and cognition (communication and social cognition) functional domains. Scores for each item range from 1 (total assistance) to 7 (complete independence). We used scores from both functional domains and the total instrument as outcomes in separate models; FIM cognition is the summation of 5 items (range: 5–35), FIM motor is the summation of 13 items (range: 13–91), and FIM total is the summation of all 18 items (range: 18–126).

**Discharge setting—**In the UDSMR database, discharge setting is coded as home, board and care, transitional living, intermediate care, skilled nursing facility, acute hospital, chronic hospital, inpatient rehabilitation, alternate level care unit, sub-acute setting, assisted living, other, or died. For this study, discharge setting was dichotomized as home versus not home.

## Covariates

Sociodemographic variables included age (years), gender, and race (white versus non-white). Clinical factors included duration from impairment onset to admission (days), length of rehabilitation stay (days), and number of comorbid conditions (range: 0–10). Admission

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

functional status was included in the models predicting discharge functional status and discharge functional status was included in the models predicting home discharge.

### Statistical Analysis

Descriptive summaries of patient characteristics and outcomes were stratified by facility volume quintiles and examined through univariate statistics. Multiple linear and logistic regression models were used to compute each patient's predicted discharge functional status and probability of home discharge, respectively. Mean predicted functional status and home discharge probabilities for all 717 facilities were calculated and plotted by volume to highlight how volume-related differences in patient casemix affect predicted outcomes. Next, we calculated the casemix-adjusted functional status and home discharge probabilities for each facility and plotted these by volume to assess if facility performance was related to volume after controlling for casemix. The casemix-adjusted functional score for each facility was calculated by dividing its actual (observed) mean FIM rating by the mean predicted (expected) FIM rating obtained in the prior step, and then multiplying this proportion by the overall mean FIM rating from all patients within the given diagnosis group. The same process (observed over expected multiplied by overall) was used to establish casemix-adjusted home discharge probabilities.[21]

Hierarchical linear modeling (HLM) was used to quantify the effects of facility volume on discharge functional status. Hierarchical generalized linear modeling (HGLM) was used to quantify the effects of facility volume on likelihood of home discharge. The hierarchical models enabled us to control for the clustering of patients within facilities and to determine the proportions of total variance in rehabilitation outcomes that were attributable to patient- and facility-level factors.[22;23] Numerical patient-level variables were grand-mean centered for inclusion in the models. Volume quintiles were dummy coded, with the middle quintile (Q3) as the reference category. Intraclass correlation coefficients (ICCs) for the empty HGLM models were estimated using the latent variable approach.[24] Separate analyses were conducted for stroke, fracture, and joint replacement. Lastly, we plotted the results of the HLM and HGLM models to show the predicted discharge FIM ratings and probabilities of home discharge by volume quintile for the typical patient within each of the three diagnosis groups. We used SPSS v.20 and HLM v.6.0 software.

Sufficient sample sizes for precision of estimates of fixed effects and variance components in multilevel analyses are a complex function of the number of higher-level units (facilities), the number of observations (patients) per facility, the ICCs for the outcomes, and the effect sizes to be detected.[25] We used the Optimal Design software[26] to assess sample sufficiency for detection of statistical differences between volume quintiles. For the HGLM analyses based on 700 facilities and a conservative scenario of 30 patients per facility, power was sufficient to detect a small (5%) difference in proportions of home discharges across facility volume quintiles at > .80. Similarly, based on 700 facilities for the HLM analyses, power to detect a small effect (d = .10) between quintile groups also exceeds .80. Due to the large sample sizes and corresponding power to detect small effects as well as the number of outcomes (analyses) included in our study, alpha was set at .01.

## RESULTS

Overall and volume-stratified patient characteristics and unadjusted outcomes for each of the three diagnosis categories are displayed in the Table. Facilities in the highest volume quintiles admit patients with lower functional status compared to lower volume facilities across all three diagnosis categories. In the stroke cohort, higher volume was also associated with a greater percentage of minority patients and a longer duration from onset to admission. Length of stay increased with volume in the stroke and fracture cohorts. Regarding

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(unadjusted) outcomes, higher volume was associated with lower mean discharge functional status in the stroke and fracture cohorts and no difference to slightly higher mean functional status among joint replacement patients. The percentage of patients discharged home decreased with increasing volume in the fracture group, whereas the stroke and joint replacement groups demonstrated slight decreases and increases, respectively.

Figures 1 displays mean estimates of discharge functional status and home discharge probabilities by diagnosis-specific volumes for each facility. The trends in the data reflect the volume-related differences in casemix depicted in the Table. Specifically, facilities with higher stroke and fracture volumes tend to admit more clinically complex patients leading to poorer predicted outcomes, whereas casemix is relatively stable across joint replacement volumes resulting in no discernible effect on predicted outcomes.

Figure 2 presents casemix-adjusted functional status and home discharge rates by diagnosis-specific volumes for each facility. Overall, the figures reveal no trend to slight increases in casemix-adjusted functional status and home discharge rates with increasing volume. There are substantial deviances in both high- and low-performing facilities throughout the volume ranges.

### HLM – discharge functional status

The ICCs demonstrated consistent rankings among diagnostic groups and varied functional status domain: FIM total (stroke = 6.1%, fracture = 9.0%, joint replacement = 15.0%); FIM motor (stroke = 5.9%, fracture = 9.3%, joint replacement = 17.9%); FIM cognition (stroke = 8.4%, fracture = 12.5%, joint replacement = 16.6%). Figure 3 shows the relationships between diagnosis-specific volume quintiles and discharge functional status from the fully-adjusted models using FIM cognition, FIM motor, and FIM total as outcome measures. The values represent the predicted functional status for a white, female patient with average demographic and clinical characteristics within each of the three diagnosis categories. Using the middle volume quintile (Q3) as the reference group, we detected small, but statistically significant ($p < .01$) associations in all three diagnosis groups. Only the highest volume quintile (Q5) differed significantly from Q3, yielding higher functional status ratings in each significant model. FIM total reached significance in all three diagnoses, FIM motor in fracture and joint replacement, and FIM cognition in fracture only.

### HGLM – discharge setting

The estimated ICCs for home discharge also varied by diagnosis group: stroke = 3.4%, fracture = 4.8%, and joint replacement = 4.9%. Figure 4 shows the relationships between diagnosis-specific volume quintiles and probability of home discharge. The values are the predicted probabilities for a white, female patient with average demographic and clinical characteristics within each of the three diagnosis categories. Again, using volume Q3 as the reference category, no significant ($p < .01$) differences in likelihood of home discharge were observed within any of the diagnostic groups.

## DISCUSSION

Our goal was to determine if the well-established volume-outcomes relationships reported in acute care extend to post-acute care. Specifically, we evaluated the effects of inpatient rehabilitation volumes on discharge functional status and return to home among patients receiving care for the three most common rehabilitation diagnoses. The multilevel models yielded small, positive, statistically significant associations between volume and discharge functional status in all three groups. Volume was not associated with probability of home discharge in any group.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Higher-volume facilities treated patients with greater average severity for the two more complex diagnosis groups: stroke and lower extremity fracture. The association between volume and casemix complexity is mixed in the acute care literature. Billingsley et al.[27] state that patients receiving colon cancer surgery from high volume providers have more favorable clinical characteristics on average compared to patients seeing lower volume surgeons. Durairaj et al.[28] showed that patients with respiratory and gastrointestinal conditions admitted to high volume intensive care units are less likely to be admitted from the emergency department, but also have higher mean severity scores. Thabut and colleagues[29] reported no relationship between annual lung transplant volume and mean patient severity scores across 61 transplant centers.

Unlike severity, there is a clear relationship between patient race/ethnicity and volume in the acute care literature. Minority patients tend to receive care in low volume hospitals.[7;30] In fact, many minority patients bypass closer, high-volume hospitals to be treated in lower-volume hospitals.[7] Little information is available regarding these patterns in inpatient rehabilitation hospitals/units, which are fewer in number than acute care hospitals; i.e. patients from both low- and high-volume hospitals are admitted to the same regional rehabilitation facility. Our data showed a distinct difference in racial proportions across stroke volume quintiles. In contrast to acute care, the average percentage of minorities in low-volume rehabilitation facilities was lower than in high-volume facilities. This finding may reflect the distribution of higher volume facilities in urban settings and geographic regions where a majority of minorities live.

Casemix-adjusted functional status and home discharge rates for each facility were calculated using the approach of Hannan and colleagues.[21] The unadjusted differences and trends in predicted outcomes that suggested poorer outcomes as facility volume increased (Table and Figure 1) were mediated by adjustments for casemix (Figure 2). It is unclear, however, whether patients with more complex conditions are more likely to experience better rehabilitation outcomes in higher volume facilities. Our data show both over- and under-performing facilities across the volume spectrum, but we did not differentiate the rehabilitation experiences of higher or lower functioning patients. Developing risk-adjustment models for assessing the quality of rehabilitative care and reporting facility-level performances is an important research area in the current healthcare reform environment.

The hierarchical models provide the most precise evidence for little-to-no relationship between diagnosis-specific volumes and rehabilitation outcomes. Volume was not associated with home discharge probability in any diagnostic group. The greatest evidence for a small volume-outcomes effect was in discharge functional status scores in the highest volume quintile (Q5) when compared to the average volume quintile (Q3). Mean Q5 scores were slightly greater than mean Q3 scores for FIM total within all three diagnoses, for FIM motor within the fracture and joint replacement groups, and for FIM cognition within the fracture group. There was no evidence of significant differences in outcomes between low volume (Q1) and average volume (Q3) facilities. It is important to note that sample size increased substantially in successive quintiles. Thus, by design, comparisons between higher-volume quintiles have greater power to detect relatively small effects. Significance aside, the largest effect was related to discharge functional status in patients with fracture, wherein the difference in mean discharge FIM total ratings between Q5 and Q3 was 3.6 points. Beninato et al.[31] report that a 22-point change in FIM total ratings is the minimal clinically important change based on attending physician assessments of clinical improvement in patients with stroke. Granger[32] describes a rule of thumb for interpreting FIM total scores in terms of projected burden of care: a 1-point increase in FIM total rating reduces the amount of daily assistance a person needs to perform essential activities by an estimated 1 to 6 minutes depending on whether he or she is on the high or low end of the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

functioning continuum. Thus, the 3.6-point difference we observed reduces the estimated daily burden of care by 4–20 minutes.

Katz et al.[33] showed that patients receiving total knee replacements from low-volume surgeons in low-volume hospitals are twice as likely to report poor functional status two years later than patients operated on by high volume surgeons in high volume hospitals. No prior rehabilitation studies were found that used multilevel modeling to examine the volume-outcomes relationship. DeJong and colleagues[34] assessed setting-specific (inpatient rehabilitation versus skilled nursing) differences in motor outcomes among patients with lower extremity joint replacement. Using a 5-level variable that combined setting and volume categories and was entered as a patient-level factor in a traditional multiple linear regression, they concluded that medium-volume (100–183 patients per year) is associated with slightly better motor outcomes than high-volume (272–347 patients per year) inpatient rehabilitation for both knee and hip replacement: 1.6 and 2.3 FIM points, respectively. Differences in analytical models notwithstanding, there are two other factors that may explain the discordance between the findings of DeJong et al. and our current results. First, they were limited to data from 11 inpatient rehabilitation facilities. Second, both their medium- and high-volume categories would be classified as high volume in our quintiles derived from the volume rankings of 717 facilities. Nevertheless, the magnitudes of the effects from both studies support the conclusion that the inpatient rehabilitation volume-outcomes relationship in joint replacement is not clinically meaningful.

The Patient Protection and Affordable Care Act (Public Law 111–148)[35] expands hospital accountability for the post-discharge recovery of Medicare beneficiaries. The current findings neither support the idea that a patient with a stroke, lower extremity fracture, or joint replacement should bypass a lower-volume local rehabilitation facility and travel to a larger-volume regional facility, nor encourage public health officials and/or third-party payers to promote further regionalization of inpatient rehabilitation services. This reasoning is based on the two most relevant rehabilitation outcomes included in this study: discharge functional status and discharge setting. Several other quality indicators are being implemented in response to provisions within the Affordable Care Act. Additional research is needed to determine if a volume benefit exists for those types of measures.

The high volume benefit in acute care has been attributed to two hypotheses: 1) practice makes perfect, wherein clinicians become more effective with more frequent encounters; and/or 2) selective referral, wherein clinicians with better reputations receive more referrals.[2;4;36;37] Moreover, the volume-outcomes relationship in acute care is more pronounced in complex high-risk medical procedures than in routine medical care.[4;28;38] Given this framework, there are several reasons why the associations between rehabilitation volume and outcomes may be less robust. First, rehabilitation is inherently multidisciplinary and multifaceted, meaning that there is no single clinician or intervention that is solely responsible for patient outcomes. Second, with less facilities available compared to acute care hospitals, there are fewer opportunities for selective referrals within a given area. Third, although rehabilitation requires more patient-clinician engagement than surgical or pharmacological interventions, it is generally less technical and procedure based.

Our approach has some limitations. We were restricted to variables in administrative datasets so facility-level factors such as resources and processes of care were not available. We did not test for cross-level interactions to assess the influence of volume on outcomes across different patient subgroups; e.g., high versus low functioning groups, racial/ethnic categories, etc. In addition, we only studied patients with the three most common diagnoses so it is plausible that most facilities achieve a volume threshold above which there is no discernible improvement in patient outcomes. The volume-outcome relationship for patients

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

with less common diagnoses who require more specialized care, such as spinal cord injury and traumatic brain injury, should be studied to test this possibility. Our study also includes several strengths. We studied patients with the three most frequent inpatient rehabilitation diagnoses and examined outcomes that are relevant to rehabilitation and meaningful to persons with substantial functional impairments. In addition, this is the first study to calculate facility-level, casemix-adjusted performance ratings and display them as a function of facility volume. This is also the first time that multilevel modeling was used to investigate the volume-outcomes relationships in post-acute settings.

In conclusion, the current findings indicate that patients with stroke, lower extremity fracture, or joint replacement need not travel to higher-volume regional rehabilitation centers; rather they can receive appropriate care in local lower-volume rehabilitation facilities. The data do not support efforts by public health officials and/or third-party payers to promote further regionalization of inpatient rehabilitation services for these conditions. However, we did identify relatively low- and high-performing facilities. In the current healthcare environment of accountable care organizations and bundled payment initiatives, performances of post-acute providers will increasingly be evaluated and this information will likely affect patient flow in the coming years as referral networks continue to be formed. From a research perspective, it is imperative that we refine the case-mix (risk) adjustment models for post-acute providers, study the volume-outcome relationship in less common diagnostic groups, and examine other modifiable facility-level factors that may improve patient-centered outcomes.

## Acknowledgments

This study was funded in part by grants from the National Institute on Disability and Rehabilitation Research (NIDRR: H133F090030 - Graham, H133G100182 - Deutsch, H133G080163 - Ottenbacher, Graham, Karmarkar) and the National Institutes of Health (NIH: R24 HD065702 - Ottenbacher, Graham, Karmarkar). The authors would like to acknowledge the insight and effort of the reviewers. Their input was invaluable in helping us clarify our findings and message.

## REFERENCES

1. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med. 1979; 301:1364–1369. [PubMed: 503167]
2. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Int Med. 2002; 137:511–520. [PubMed: 12230353]
3. Holt PJ, Poloniecki JD, Gerrard D, Loftus IM, Thompson MM. Meta-analysis and systematic review of the relationship between volume and outcome in abdominal aortic aneurysm surgery. Br J Surg. 2007; 94:395–403. [PubMed: 17380547]
4. Nuttall M, van der MJ, Phillips N, et al. A systematic review and critique of the literature relating hospital or surgeon volume to health outcomes for 3 urological cancer procedures. J Urol. 2004; 176:2145–2152. [PubMed: 15538220]
5. Chassin MR. Improving the quality of health care: what strategy works? Bull N Y Acad Med. 1996; 73:81–91. [PubMed: 8804741]
6. Epstein AM. Volume and outcome--it is time to move ahead (Editorial). N Engl J Med. 2002; 346:1161–1164. [PubMed: 11948278]
7. Losina E, Wright EA, Kessler CL. Neighborhoods matter: use of hospitals with worse outcomes following total knee replacement by patients from vulnerable populations. Arch Intern Med. 2007; 167:182–187. [PubMed: 17242320]
8. Birkmeyer JD, Finlayson EV, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. Surgery. 2001; 130:415–422. [PubMed: 11562662]
9. Shahian DM, Normand SL. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg. 2003; 75:1048–1058. [PubMed: 12645752]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript


10. Kizer KW. The volume-outcome conundrum (Editorial). N Engl J Med. 2003; 349:2159–2161. [PubMed: 14645645]
11. Agency for Healthcare Research and Quality. Guide to Inpatient Quality Indicators: Quality of Care in Hospitals - Volume, Mortality, and Utilizaton. 2007 Dec 3.
12. Li Y, Cai X, Yin J, Glance LG, Mukamel DB. Is higher volume of postacute care patients associated with a lower rehospitalization rate in skilled nursing facilities? Med Care Res Rev. 2012; 69:103–118. [PubMed: 21810798]
13. Mallinson TR, Manheim LM, Almagor O, Demark HM, Heinemann AW. Trends in the supply of inpatient rehabilitation facilities services: 1996 to 2004. Arch Phys Med Rehabil. 2008; 89:2066–2079. [PubMed: 18996234]
14. US Department of Health and Human Services. Federal Register. Vol. Vol. 73. Washington, DC: 2008 Aug 8.
15. Langhorne P, Duncan P. Does the organization of postacute stroke care really matter? Stroke. 2001; 32:268–274. [PubMed: 11136947]
16. Medicare Payment Advisory Commission. Washington, DC: 2011. Report to the Congress: Medicare Payment Policy.
17. UDSMR. Uniform Data System for Medical Rehabilitation. 2011 http://udsmr.org.
18. UB Foundation Activities. The Inpatient Rehabilitation Facility - Patient Assessment Instrument (IRF-PAI) Training Manual. 2004. http://www.cms.hhs.gov/InpatientRehabFacPPS/downloads/irfpaimanual040104.pdf Available from: Centers for Medicare and Medicaid Services
19. Granger CV, Cotter AC, Hamilton BB, Fiedler RC. Functional assessment scales: a study of persons after stroke. Arch Phys Med Rehabil. 1993; 74:133–138. [PubMed: 8431095]
20. Ottenbacher KJ, Hsu Y, Granger CV, Fiedler RC. The reliability of the functional independence measure: a quantitative review. Arch Phys Med Rehabil. 1996; 77:1226–1232. [PubMed: 8976303]
21. Hannan EL, Racz MJ, Jollis JG, Peterson ED. Using Medicare claims data to assess provider quality for CABG surgery: does it work well enough? Health Serv Res. 1997; 31:659–678. [PubMed: 9018210]
22. Raudenbush, SW.; Bryk, AS. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd ed.. Thousand Oaks: CA: Sage Publications; 2002.
23. Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. Ann Int Med. 2003; 139:658–665. [PubMed: 14568854]
24. Goldstein H, Browne W, Rasbash J. Partitioning variation in multilevel models. Understanding Statistics. 2002:223–231.
25. Maas CJ, Hox JJ. Sufficient sample sizes for multilevel modeling. Methodology: Eur J Res Methods Behav Soc Sci. 2005; 1:86–92.
26. Spybrook, J. Power, sample size, and design. In: O'Connel, AA.; McCoach, DB., editors. Multilevel Modeling of Educational Data. Charlotte, NC: Information Age Publishing; 2008. p. 273-311.
27. Billingsley KG, Morris AM, Dominitz JA. Surgeon and hospital characteristics as predictors of major adverse outcomes following colon cancer surgery: understanding the volume-outcome relationship. Arch Surg. 2007; 142:23–31. [PubMed: 17224497]
28. Durairaj L, Torner JC, Chrischilles EA, Vaughan Sarrazin MS, Yankey J, Rosenthal GE. Hospital volume-outcome relationships among medical admissions to ICUs. Chest. 2005; 128:1682–1689. [PubMed: 16162775]
29. Thabut G, Christie JD, Kremers WK, Fournier M, Halpern SD. Survival differences following lung transplantation among US transplant centers. JAMA. 2010; 304:53–60. [PubMed: 20606149]
30. Epstein AJ, Gray BH, Schlesinger M. Racial and ethnic differences in the use of high-volume hospitals and surgeons. Arch Surg. 2010; 145:179–186. [PubMed: 20157087]
31. Beninato M, Gill-Body KM, Salles S, Stark PC, Black-Schaffer RM, Stein J. Determination of the minimal clinically important difference in the FIM instrument in patients with stroke. Arch Phys Med Rehabil. 2006; 87:32–39. [PubMed: 16401435]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

32. Granger CV. The emerging science of functional assessment: our tool for outcomes analysis. Arch Phys Med Rehabil. 1998; 79:235–240. [PubMed: 9523772]

33. Katz JN, Mahomed NN, Baron JA. Association of hospital and surgeon procedure volume with patient-centered outcomes of total knee replacement in a population-based cohort of patients age 65 years and older. Arthritis Rheum. 2007; 56:568–574. [PubMed: 17265491]

34. DeJong G, Horn SD, Smout RJ, Tian W, Putman K, Gassaway J. Joint replacement rehabilitation outcomes on discharge from skilled nursing facilities and inpatient rehabilitation facilities. Arch Phys Med Rehabil. 2009; 90:1284–1296. [PubMed: 19651262]

35. 111th Congress. Patient Protection and Affordable Care Act (Public Law 111–148). 2010 May 1.

36. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res. 1987; 22:157–182. [PubMed: 3112042]

37. Flood AB, Scott WR, Ewy W. Does practice make perfect? Part I: The relation between hospital volume and outcomes for selected diagnostic categories. Med Care. 1984; 22:98–114. [PubMed: 6700280]

38. Birkmeyer JD, Siewers AE, Finlayson EV. Hospital volume and surgical mortality in the United States. N Engl J Med. 2002; 346:1128–1137. [PubMed: 11948273]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

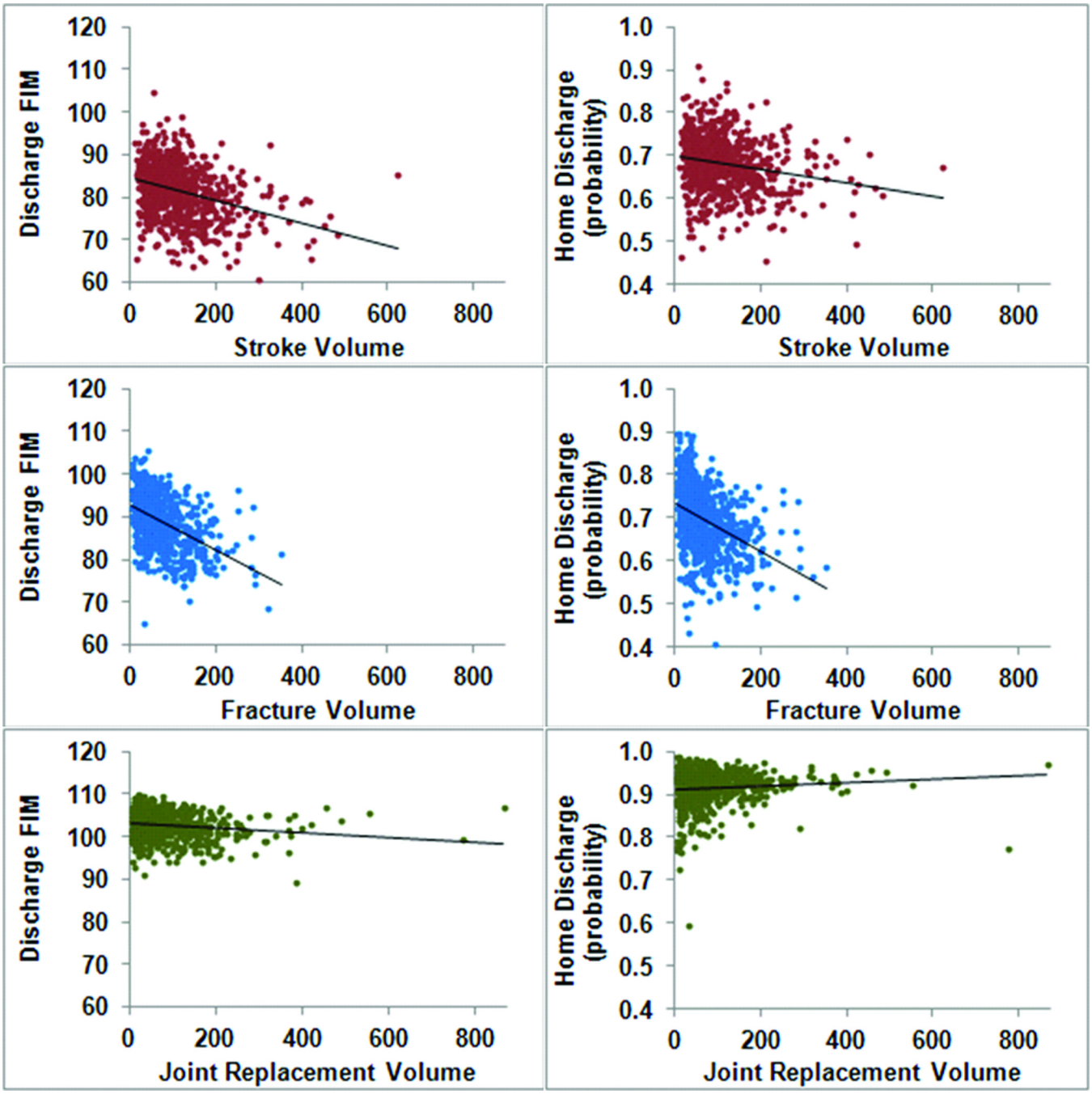


**Figure 1.**

Mean predicted discharge FIM ratings and home discharge probabilities for all facilities (N = 717 stroke; N = 713 fracture; N = 712 joint replacement) by diagnosis-specific volumes. Multiple linear regression was used to compute each patient's predicted discharge FIM rating, from which the mean FIM rating for each facility was calculated. Multiple logistic regression was used to compute each patient's probability of home discharge, from which the mean home discharge probability for each facility was calculated. The scores reflect the expected mean values for each facility when applying its specific case-mix to the average effects derived from the entire diagnosis-specific population.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

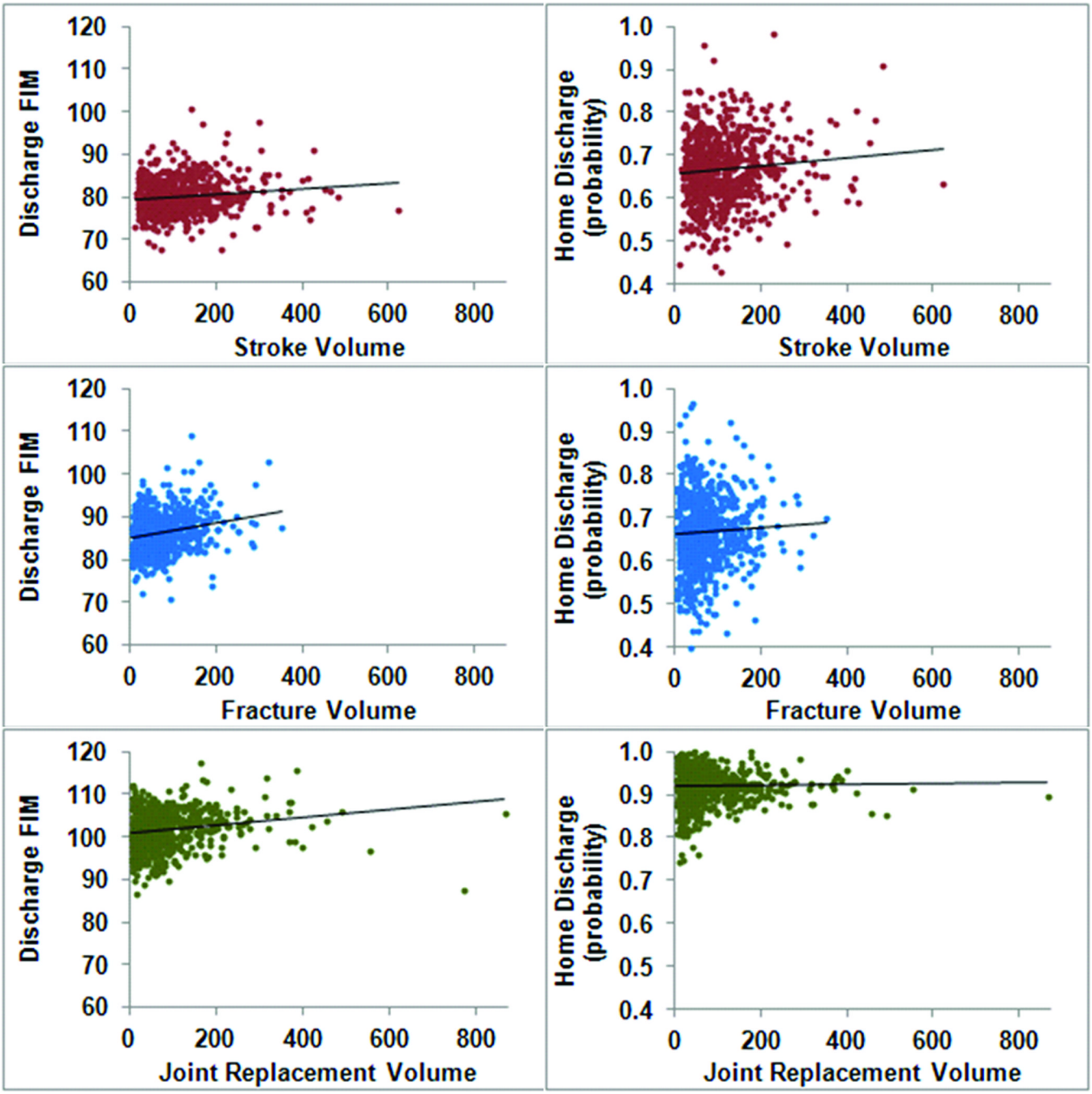


**Figure 2.**

Casemix-adjusted mean discharge FIM ratings and home discharge probabilities for all facilities (N = 717 stroke; N = 713 fracture; N = 712 joint replacement) by diagnosis-specific volumes. Values were calculated by taking a facility's actual (observed) mean values and dividing by the predicted (expected) values for that facility (displayed in Figure 1), and multiplying those ratios by the grand mean FIM rating and home discharge probability, respectively, within each diagnosis category. This is a form of indirect standardization, wherein the scores reflect the mean outcomes experienced in each facility relative to the outcomes experienced by the entire diagnosis-specific population with a similar case-mix.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



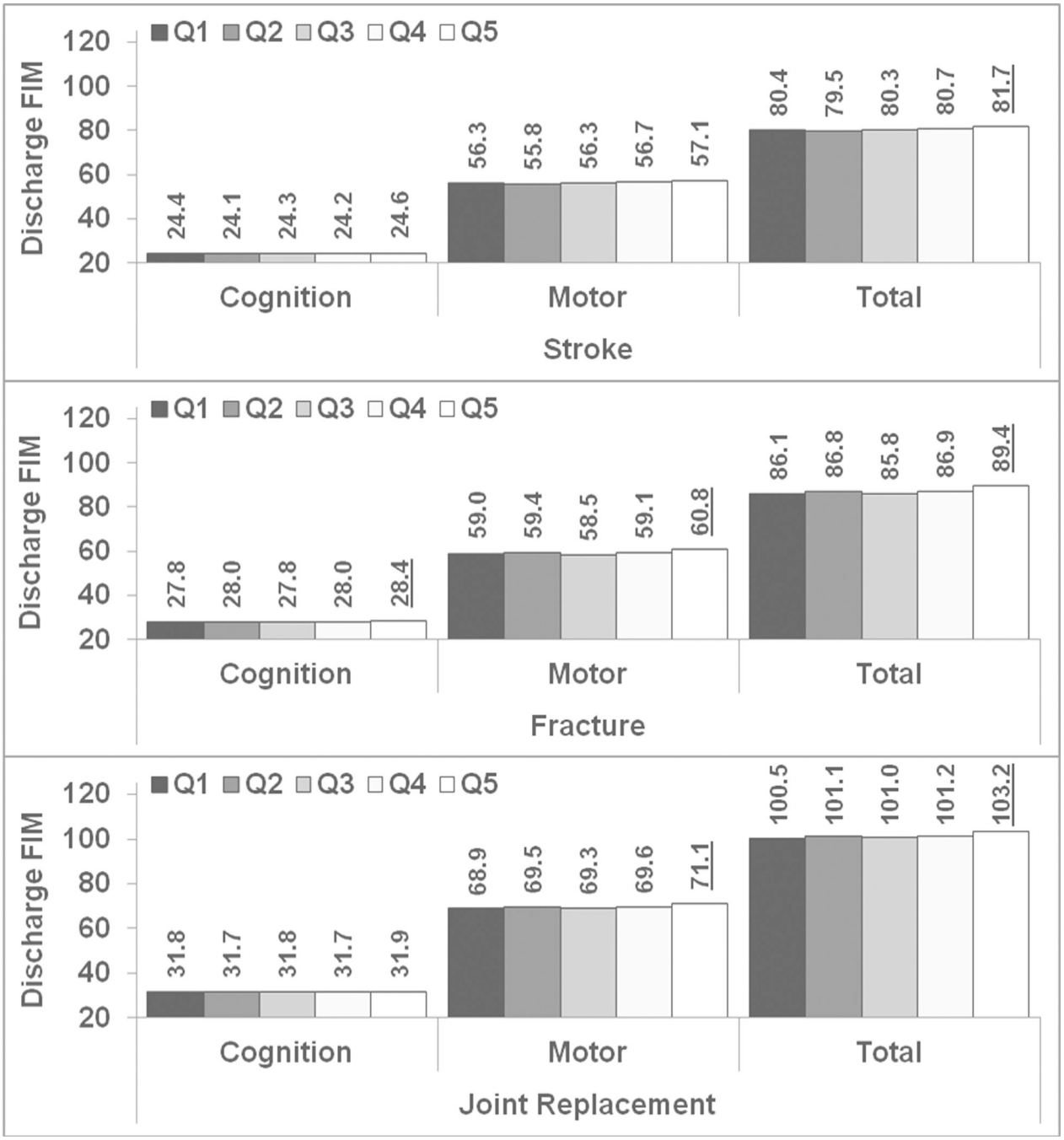


**Figure 3.**

Predicted discharge cognition, motor, and total functional status ratings by diagnosis-specific volume quintiles. The values were obtained from hierarchical linear models and represent the predicted FIM ratings for white, female patients with average clinical characteristics (age, comorbidity burden, duration to admission, admission functional status, and length of stay) within each of the three diagnosis categories. Sample sizes increase across quintiles and vary by diagnosis: stroke 12,867 (Q1) – 83,567 (Q5), fracture 5,609 (Q1) – 59,511 (Q5), joint replacement 4,326 (Q1) – 79,831 (Q5), see Table for more details. Reference category = Q3 for all tests of significance. Underlined value labels indicate significant ($p < .01$) difference compared to Q3.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

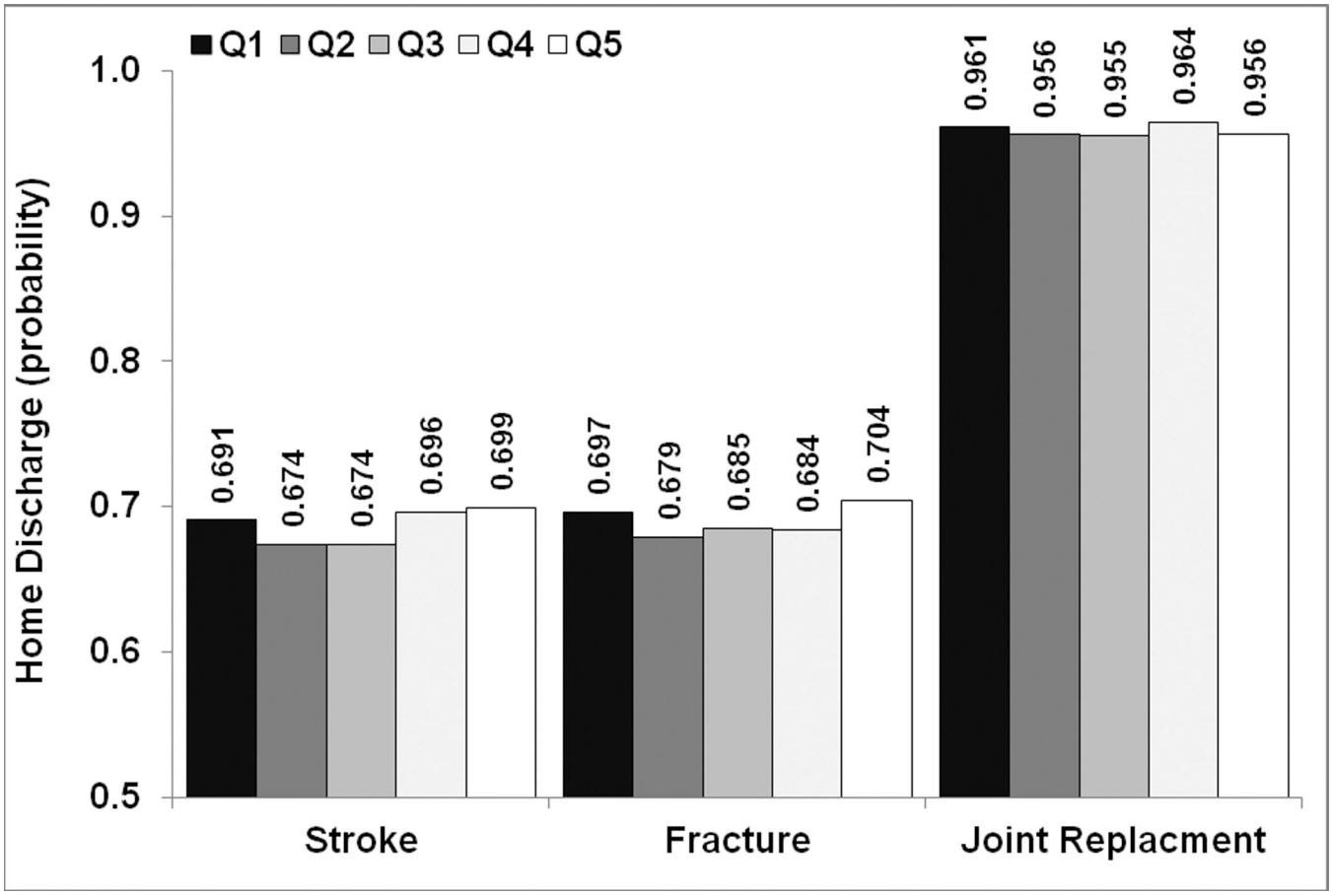


**Figure 4.**

Predicted home discharge probability by diagnosis-specific volume quintiles. The values were obtained from hierarchical generalized linear models and represent the predicted probabilities for white, female patients with average clinical characteristics (age, comorbidity burden, duration to admission, admission functional status, and length of stay) within each of the three diagnosis categories. Sample sizes increase across quintiles and vary by diagnosis: stroke 12,867 (Q1) – 83,567 (Q5), fracture 5,609 (Q1) – 59,511 (Q5), joint replacement 4,326 (Q1) – 79,831 (Q5), see Table for more details. Reference category = Q3 for all tests of significance, none of which reached $p < .01$.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

## Table

Patient characteristics and unadjusted outcomes for patients with stroke, lower extremity fracture and lower extremity replacement: total samples and by diagnosis-specific volume quintiles.

| | | Facility Volume Quintiles [range] | | | | |
|---|---|---|---|---|---|---|
| **Stroke** | **Total** | **Q1 [< 52]** | **Q2 [52–79]** | **Q3 [80–112]** | **Q4 [113–164]** | **Q5 [> 164]** |
| Facilities, N | 717 | 140 | 148 | 142 | 143 | 144 |
| Patients, N | 202,423 | 12,867 | 23,621 | 34,542 | 47,826 | 83,567 |
| Annual volume | 112.9 (78.5) | 36.2 (10.5) | 63.0 (8.1) | 95.4 (10.0) | 133.1 (14.3) | 236.0 (77.9) |
| Age, yrs | 71.5 (11.9) | 73.2 (11.6) | 72.5 (11.7) | 71.7 (12.0) | 71.3 (12.0) | 71.0 (12.0) |
| Women | 51.8% | 53.8% | 51.7% | 52.4% | 51.3% | 51.5% |
| White | 73.2% | 83.9% | 79.3% | 70.0% | 72.8% | 71.3% |
| Comorbid sum, 0–10 | 8.0 (2.5) | 8.0 (2.5) | 8.0 (2.5) | 7.8 (2.6) | 7.9 (2.5) | 8.1 (2.3) |
| Onset to admit, days | 8.5 (8.3) | 7.5 (7.7) | 7.7 (7.5) | 8.1 (8.1) | 8.3 (8.1) | 9.1 (8.8) |
| Admission FIM | 55.1 (19.1) | 58.7 (19.3) | 58.5 (18.7) | 57.6 (18.9) | 55.8 (19.1) | 52.1 (18.9) |
| Discharge FIM | 80.1 (23.7) | 82.8 (23.7) | 82.1 (23.3) | 82.0 (23.2) | 80.7 (23.5) | 78.1 (24.0) |
| FIM change | 25.0 (14.8) | 24.0 (14.8) | 23.6 (14.7) | 24.3 (14.6) | 24.9 (14.8) | 25.9 (15.0) |
| Length of stay, days | 16.7 (8.9) | 15.3 (8.6) | 15.8 (8.4) | 16.1 (8.8) | 16.3 (8.7) | 17.6 (9.1) |
| Discharged home | 67.4% | 68.0% | 67.1% | 68.1% | 68.4% | 66.5% |
| **Fracture** | **Total** | **Q1 [< 27]** | **Q2 [27–44]** | **Q3 [45–66]** | **Q4 [67–104]** | **Q5 [> 104]** |
| Facilities, N | 713 | 147 | 145 | 140 | 141 | 141 |
| Patients, N | 132,194 | 5,609 | 14,128 | 20,778 | 32,168 | 59,511 |
| Annual volume | 68.8 (53.9) | 14.8 (7.8) | 36.5 (5.2) | 55.3 (6.5) | 84.0 (10.9) | 157.3 (46.5) |
| Age, yrs | 78.4 (10.4) | 76.3 (11.7) | 77.8 (10.7) | 78.3 (10.3) | 78.4 (10.4) | 78.9 (10.1) |
| Women | 72.1% | 67.6% | 70.5% | 71.6% | 72.4% | 72.9% |
| White | 88.0% | 86.7% | 86.2% | 89.7% | 87.9% | 88.0% |
| Comorbid sum, 0–10 | 7.8 (2.5) | 7.4 (2.7) | 7.6 (2.6) | 7.5 (2.7) | 7.9 (2.5) | 8.0 (2.3) |
| Onset to admit, days | 6.3 (5.4) | 7.0 (6.4) | 6.3 (5.4) | 6.1 (5.1) | 6.1 (5.0) | 6.4 (5.5) |
| Admission FIM | 60.5 (15.5) | 66.4 (14.4) | 65.8 (14.8) | 64.1 (14.3) | 61.2 (15.2) | 57.0 (15.5) |
| Discharge FIM | 86.9 (19.1) | 90.6 (17.7) | 90.4 (18.2) | 88.4 (18.3) | 86.6 (19.1) | 85.4 (19.5) |
| FIM change | 26.5 (13.4) | 24.2 (12.5) | 24.6 (12.4) | 24.4 (12.7) | 25.4 (13.1) | 28.4 (13.9) |

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

| | | Facility Volume Quintiles [range] | | | | |
|---|---|---|---|---|---|---|
| Length of stay, days | 13.4 (5.3) | 12.1 (5.7) | 12.0 (5.0) | 12.7 (5.2) | 13.5 (5.2) | 14.1 (5.3) |
| Discharged home | 67.3% | 73.1% | 71.6% | 68.7% | 66.3% | 65.9% |
| **Joint Replacement** | **Total** | **Q1 [< 20]** | **Q2 [20–37]** | **Q3 [38–63]** | **Q4 [64–118]** | **Q5 [> 118]** |
| Facilities, N | 712 | 147 | 142 | 140 | 142 | 141 |
| Patients, N | 148,068 | 4,326 | 10,985 | 18,795 | 34,131 | 79,831 |
| Annual volume | 75.5 (85.3) | 10.9 (5.5) | 28.0 (4.9) | 49.5 (7.3) | 86.2 (15.1) | 206.0 (109.6) |
| Age, yrs | 70.5 (10.4) | 70.0 (11.1) | 70.9 (10.8) | 70.9 (10.7) | 70.8 (10.3) | 70.2 (10.2) |
| Women | 68.3% | 65.9% | 67.7% | 68.7% | 69.2% | 68.1% |
| White | 83.4% | 81.0% | 84.7% | 82.3% | 82.8% | 83.8% |
| Comorbid sum, 0–10 | 6.9 (2.7) | 6.7 (2.8) | 6.9 (2.7) | 6.8 (2.7) | 7.0 (2.7) | 7.0 (2.6) |
| Onset to admit, days | 4.3 (3.6) | 5.0 (4.9) | 4.6 (4.1) | 4.6 (3.9) | 4.3 (3.8) | 4.2 (3.4) |
| Admission FIM | 73.2 (12.7) | 75.6 (12.9) | 75.7 (12.3) | 74.7 (12.4) | 74.5 (12.4) | 71.8 (12.7) |
| Discharge FIM | 102.2 (11.9) | 101.1 (12.9) | 102.0 (12.0) | 101.6 (12.1) | 101.7 (11.8) | 102.6 (11.7) |
| FIM change | 29.0 (11.7) | 25.6 (11.0) | 26.3 (10.7) | 26.9 (11.0) | 27.3 (11.1) | 30.8 (12.0) |
| Length of stay, days | 9.2 (3.9) | 9.2 (4.4) | 8.9 (4.1) | 9.3 (4.1) | 9.0 (3.9) | 9.3 (3.8) |
| Discharged home | 92.0% | 91.2% | 91.3% | 91.1% | 92.6% | 92.2% |

Values reported are mean (sd) or column percentages. Differences between quintile groups tested with one-way ANOVA or chi-square tests: $p < .01$ for all comparisons. Volume ranges defining quintiles are noted in [brackets] within the diagnostic-specific column-header rows.