The Journal of Arthroplasty Vol. 00 No. 0 2009

All-Patient Refined Diagnosis-Related Groups in Primary Arthroplasty

Carlos J. Lavernia, MD,*† Artit Laoruengthana, MD,† Juan S. Contreras, MD,† and Mark D. Rossi, PhD, PT, CSCS‡

Abstract: Our objective was to determine if the All-Patient Refined Diagnosis-Related Groups (APR-DRGs) and other comorbidity scores correlate with pain level, functional abilities, and hospital cost after primary total joint arthroplasty (TJA). Three hundred three patients having TJA were evaluated with average follow-up of 21 months. Western Ontario and McMaster Universities osteoarthritis index, Short-Form 36, and Quality of Well-Being index were administered before and after surgery. The APR-DRG subclassification including severity of illness (SOI) subclass scores and risk of mortality (ROM), Charlson index, American Society of Anesthesiologist (ASA), Charnley score, length of stay, and hospital costs were reported. Patients in a higher SOI and ROM subclasses had a statistically significant decrease in functional outcomes scores, longer length of stay, and greater hospitals costs than those in lower subclasses. However, correlations of comorbidity categories with outcome scores were poor. The APR-DRG classification helps identify those individuals with worse function and is specially suited in identifying those patients who incur a higher hospital cost. **Keywords:** severity of illness, risk of mortality, primary arthroplasty, charlson index, charnley score.

Currently, rising health care costs in conjunction with reduced hospital reimbursement have brought quality reporting and pay for performance programs to the forefront of health policy decision making. These programs are initiatives that link financial incentives to specific performance measures with the goal of improving quality and efficiency in the health care systems [1]. These data must be risk adjusted to create a level playing field in which differences in quality care can be separated from differences in outcome attributable to severity of illness (SOI). All Patient Refined-Diagnosis Related Groups (APR-DRGs) is a software algorithm that measures SOI and likelihood of dying. It is based on the amount and severity of medical conditions present on admission. It has become one of the most common systems for SOI and risk of mortality (ROM) adjustment. In 2006, The Centers for Medicare and Medicaid Services' recommended using APR-DRGs as the primary predictor of resource use [2]. The ability of these indices to serve as a SOI adjustor in primary lower extremity joint arthroplasty remains controversial. The purpose of this study was to assess the relationship of comorbid indices and the APR-DRGs to patient outcomes in primary arthroplasty.

Materials and Methods

A prospective study was performed on 303 procedures from a single total joint arthroplasty center. Patients were included if they received a primary hip or knee arthroplasty from June 2003 to December 2005. Exclusion criteria included an infected joint, simultaneous bilateral arthroplasty, or staged bilateral arthroplasty in the same admission. All surgeries were performed by the senior author. Patients had an average follow-up of 21 (SE ± .51) months. Patient demographics are presented in Table 1. Clinical factors included in the analysis were principal diagnosis, APR-DRG classification, Charnley score, Charlson index, and American Society of Anesthesiologist (ASA) score. All demographic data and comorbidity scales were collected from medical records including anesthesia and medical consultation reports. The SOI and ROM subclasses for the APR-DRG were obtained from the hospital financial services, which were calculated using proprietary software that is based on all diagnoses covering the discharge period.

*From the *Orthopedic Institute at Mercy Hospital, Miami, Florida; †Arthritis Surgery Research Foundation Inc, Miami, Florida; and ‡Department of Physical Therapy, Florida International University, University Park, Miami, Florida.*

Submitted January 23, 2009; accepted March 15, 2009.

Ethical Board Review statement: The authors certify that a local institutional review board has approved this investigation and informed consent was obtained from each participant.

No benefits or funds were received in support of this study.

Reprint requests: Carlos J. Lavernia, MD, Orthopedic Institute at Mercy, 3659 S. Miami Ave Ste 4008, Miami, FL 33133.

0883-5403/09/0000-0000$36.00/0
doi:10.1016/j.arth.2009.03.008

Table 1. Subject Characteristics Before Surgery

	THA		TKA		Cohort	
	Male	Female	Male	Female	Male	Female
Frequency	33	57	58	155	91	212
Mean age	68.02	68.83	74.20	72.94	73.78	73.75

The APR-DRG SOI determines the extent of system breakdown or organ dysfunction, whereas the ROM determines the likelihood of dying. Both subclasses have 4 levels ranging from 1 to 4. A higher number indicates multiple, serious diseases, and associated interaction. There are 4 levels of the SOI subclass. Because we had only one subject in the fourth level, we grouped that subject into the third level, thus, having 3 groups. Risk of mortality also had 4 levels; however, in our study, there was only one patient in the fourth level and 8 patients in the third level; we then grouped the patient in the fourth level and the 8 patients in the third level into the second level, yielding 2 groups for the analysis.

The Charnley score classifies joint disease into category of involvement A (single) and B (bilateral), and category C was described as multiple joints and systemic diseases such as inflammatory arthritis and cardiovascular disability that affect physical function [3]. The Charlson index covers 19 categories of disease [4]. In our study, we divided the Charlson index score into 3 groups as follows: low (score = 0), medium (score = 1-2), and severe (score >2). The American Society of Anesthesiologist score is a preoperative rating assigned to each patient and is a measure of the patient's general health status and coexisting conditions. Scores range from 1, representing a healthy person, to 5, representing a patient not expected to survive longer than 24 hours. For our study, we grouped the ASA from 5 score levels into 2 groups (<3 and ≥3) because only 3 patients were in level I, 3 in level IV, and 2 patients in level V.

To assess quality of life and functional outcomes before and yearly after surgery, we used the Western Ontario and McMaster Universities osteoarthritis index (WOMAC), Short-form 36 (SF-36), and Quality of Well-Being index. We also compared length of stay and hospital costs between different levels of SOI subclass and ROM subclasses, as well as ASA score, Charnley score, and Charlson index score. We grouped hospital cost into gross revenue, fixed cost, and variable cost. *Gross revenue* was defined as the total charges. The variable costs are those costs related to the amount of supplies or labor for a procedure charge code (as opposed to overhead). Fixed costs are the overhead costs allocated proportionately to the departmental charge codes.

Analysis

For the Complexity subclass, Charnley, and Charlson indices we used a 2 × 3 analysis of variance (ANOVA) to assess for differences between groups before and after surgery for each outcome measure. The first factor was time (pre-post), and the second factor was group (3 subclasses). For risk of mortality and ASA, we used a 2 × 2 ANOVA to assess for differences between groups before and after surgery. For both ANOVAs, we used the measure command in SPSS (SPSS Inc, Chicago, Ill) to define all dependent measures in one model thus using a doubly multivariate approach. Follow-up tests were used as necessary when significant interactions or main effects were found. Between-group comparisons for any significant main effect of group were reduced by using the least significant difference. Age and sex were used as covariates in all analyses.

We used a 1-way ANOVA to assess for differences between groups (level of subclass) for the complexity subclass, Charnley, and Charlson indices for dependent measures of gross revenue, net revenue, fixed cost, and variable cost. Any significant omnibus test was followed up with post hoc tests using Tukey test. For the risk of mortality and ASA, we used independent t tests to assess for differences between groups for dependent measures of gross revenue, net revenue, fixed cost, and variable cost. A P value less than .05 was considered significant for all tests. We used a Spearman ϱ correlation coefficient to determine the association between comorbidity classifications and outcome measures. Knowing that most patient's functional abilities improve from before to after surgery, we used the change score of all outcome measures as the dependent variables in which to evaluate the relationships between outcomes and comorbidity classifications.

Results

Severity of Illness

For the SOI subclass, the sample sizes were 64 for group 1, 158 for group 2, and 81 for group 3. The global ANOVA indicated that there was a significant interaction for the WOMAC physical function score (P = .009; power, 0.80) and the WOMAC total score (P = .013; power, 0.76). Severity of illness group 3 reported significantly less physical function than those individuals in SOI group 1 ($P \leq .0001$) and group 2 ($P \leq .0001$) before surgery. For the WOMAC total score, patients in SOI group 3 reported significantly worse scores compared to those in group 1 (P = .001) and group 2 (P = .001) before surgery. There were no significant interactions for any other dependent measure; however, there was a main effect of time for all outcome measures ($P \leq .0001$). Regardless of group, all scores improved postoperatively.

There was also a main effect of SOI group for the SF-36 physical function score (P = .011; power, 0.78), SF-36 physical component score (P = .009; power, 0.79), WOMAC function score (P = .02; power, 0.72), and WOMAC total score (P = .025; power, 0.68). At follow-up, patients in SOI group 3 scored significantly worse than those group 2 for the SF-36 physical function (P =

Table 2. Differences in SOI groups for Hospital Cost and Length of Stay

Measure	SOI Group	Mean	SE ±	*P*
Gross revenue (US$)	1	$61 539.14	$929.80	.01
	2	$65 856.72	$1105.09	
	3	$71 025.47	$3721.97	
Variable cost (US$)	1	$10 589.14	$200.12	.03
	2	$11 001.57	$179.75	
	3	$11 826.31	$480.36	
Fixed cost (US$)	1	$4718.27	$113.19	≤.0001
	2	$5148.82	$146.06	
	3	$6220.89	$444.29	
Length of stay (d)	1	3.72	0.123	≤.0001
	2	4.15	0.136	
	3	4.82	0.227	

.003), SF-36 physical component (*P* = .002), WOMAC function (*P* = .009), and WOMAC total (*P* = .011). Moreover, those patients in group 3 scored significantly worse than those in group 1 for WOMAC function (*P* = .02) and WOMAC total (*P* = .03). Examining hospital costs, multiple comparisons indicated that when compared to SOI group 3; group1 had significantly less gross revenue (*P* = .01), variable cost (*P* = .03), fixed cost (*P* ≤ .0001), and length of stay (*P* ≤ .0001) (Table 2). Moreover, SOI group 2 had significantly less fixed cost than group 3 (*P* = .004) and less length of stay (*P* = .01) than those in group 3.

Risk of Mortality

The sample size for the ROM subclass was 210 for ROM group 1 and 93 for group 2. The omnibus test for ROM indicated that there was no time by group interaction for any dependent variable. There was a main effect of time for all SF-36 and WOMAC. All patients significantly improved postoperatively (*P* ≤ .0001). Patients in group 2 had worse scores than those in group 1 for the SF-36 physical function (*P* = .002), SF-36 general health (*P* = .019), WOMAC function (*P* = .019), and WOMAC total (*P* = .034). In addition, patients in ROM group 2 had greater gross revenue (*P* = .009), variable cost (*P* = .02), fixed cost (*P* = .01), and greater length of stay (*P* = .002) than patients in ROM group 1 (Table 3).

Table 3. Differences in ROM Groups for Hospital Cost and Length of Stay

Measure	ROM Group	Mean	SE ±	*P*
Gross revenue (US$)	1	$63 460.04	$719.64	.009
	2	$72 666.57	$3370.28	
Variable cost (US$)	1	$10 790.36	$131.17	.02
	2	$11 899.34	$440.62	
Fixed cost (US$)	1	$4495.17	$106.37	.01
	2	$6022.66	$374.14	
Length of stay (d)	1	4.01	0.103	.002
	2	4.76	0.215	

Table 4. Correlation Between Comorbidity Scores And Outcome Measures

	SOI	ROM	Charlson Score	Charnley Score	ASA
Δ QWB	0.081	0.043	−0.055	0.068	0.036
Δ $SF\text{-}36_{PF}$	−0.039	−0.096	−0.013	0.056	0.028
Δ $SF\text{-}36_{BP}$	−0.038	−0.031	0.008	0.005	0.066
Δ $SF\text{-}36_{GH}$	0.077	−0.009	0.055	0.000	0.120
Δ $SF\text{-}36_{SF}$	0.124	0.060	0.017	−0.039	0.138
Δ $SF\text{-}36_{PC}$	−0.048	−0.052	−0.028	0.084	0.065
Δ $WOMAC_F$	0.201	−0.006	0.012	0.052	0.081
Δ $WOMAC_P$	0.133	0.003	0.090	0.042	0.117
Δ $WOMAC_T$	0.205	−0.009	0.021	0.040	0.074

QWB indicates Quality of Well Being index; $SF\text{-}36_{PF}$, physical functioning; $SF\text{-}36_{BP}$, bodily pain; $SF\text{-}36_{GH}$, general health; $SF\text{-}36_{SF}$, social functioning; $SF\text{-}36_{PC}$, physical component summary; $WOMAC_F$, WOMAC function; $WOMAC_P$, WOMAC pain; $WOMAC_T$, WOMAC total.

Charnley Score

The sample size for the Charnley comorbidity scale was 149 for group 1 (category A), 62 for group 2 (category B), and 92 for group 3 (category C). The omnibus test indicated that there was a time by group interaction for the SF-36 physical function (*P* = .001; power, 0.91) score. Follow-up tests indicated that postsurgery, those in group 3, had worse scores than individuals in group 2 (*P* = .03). There was a main effect of time for all outcome measures as all categories improved postoperatively (range, *P* ≤ .0001-.017). There was no difference between groups for any dependent measure. There was also no difference between groups for any cost measure or length of stay.

Charlson Index

The sample size for the Charlson index was 73 for group 1 (low), 131 for group 2 (medium), and 99 for group 3 (high). The analysis of variance indicated that there was no time by group interaction for any dependent variable. There was a main effect of time for all outcome measures; regardless of group, all scores improved significantly postoperatively (range, *P* ≤ .0001-.007). There was no main effect of group for any dependent measure. Pertaining to cost, there was no difference between groups for any outcome cost measure or length of stay.

American Society of Anesthesiologists

The global ANOVA indicated an interaction for the SF-36 physical function (*P* = .003). Follow-up test indicated that group 2 was significantly worse than group 1 before surgery (*P* = .002). However, these differences did not prevail after surgery. All patients demonstrated improvement postoperatively for Quality of Well Being index (*P* = .002), SF-36 physical function (*P* = .015), SF-36 physical component (*P* = .001), WOMAC pain (*P* ≤ .0001), WOMAC function (*P* ≤ .0001), and WOMAC total (*P* ≤ .0001) scores. There was a main effect of group for the WOMAC physical function ($P = .04$; power, .60) and WOMAC total ($P = .03$; power, .60). Although there was no difference between groups for all cost measures, those patients in group 2 had longer length of stay (4.6 days) when compared with those patients in group 1 (3.9 days) ($P = .004$).

Correlations

The association of comorbidity categories with outcome Δ scores were poor, with the SOI subclass having better correlations with WOMAC change scores than all other comorbidity categories (Table 4).

Discussion

Value purchasing is one of the important mechanisms to reduce health care costs without rationing. Pay for performance and transparency with outcomes have also been touted as a means of improving the way in which health care is delivered in the United States. They provide a mechanism for financial incentives when better outcomes are achieved [5]. Although these programs have continued to grow, important questions have been raised as to how performance measures are defined and how to "level the playing field" for those physicians caring for the sicker patients. A multiplicity of research and clinical indices to assess SOI exist. The APR-DRG has been reported to be a valuable predictor of resource use in total joint arthroplasty [2].

In our study, the APR-DRGs were better predictors of costs than outcomes (Tables 2 and 3). We did find that those individuals who were in a higher SOI subclass reported greater physical dysfunction when compared to those in a lower subclass. The ROM subclass was a predictor of low scores in physical health outcome measures (SF-36 physical function [$P = .002$], and WOMAC function [$P = .019$]). Individuals in all SOI and ROM groups improved over time after the surgery and the Δ scores showed poor correlations with the APR-DRG indices. Our data demonstrate a poor correlation between outcomes and time and comorbid categorization using all scales (Table 4).

All outcome measures improved postoperatively for all of the Charnley groups. However, there were no differences between groups for any outcome measure or any differences between Charnley score groups for length of stay or cost. This is in disagreement with the data published by Wasielewski et al [6] who reported a significant association between individuals who underwent total knee arthroplasty with higher Charnley scores and increased length of stay and hospital cost. Our data demonstrate no correlation between the Charlson index and all outcome measures and hospital costs. This is in disagreement with Ethgen et al [7] who reported that those individuals with severe comorbid status stratified by Charlson index had associated worse function scores after knee arthroplasty.

The ASA score was a predictor of low preintervention scores in physical outcome measures. Moreover, those with higher scores and greater health risk had significantly longer length of stay (4.6 vs 3.9 days). This is in disagreement with Wasielewski et al [6] who reported that in a small cohort of patients undergoing total knee arthroplasty, the ASA had no correlation with outcome, length of stay, or complication rate. However, the authors did report that ASA classification could predict hospital costs. In our cohort, there were no correlations between the ASA and the financial parameters. Cullen et al [8] evaluating patients undergoing primary total hip arthroplasty found that high ASA score was not predictive for length of stay but was predictive of complication rate after surgery. Similar to our study, Greenfield et al [9] reported that a higher ASA score was associated with longer length of stay.

A limitation of our study was that we did not have enough patients in our study group categorized into the extreme or worse SOI and ROM groups; thus, our sample may be not representative for the whole population. Because of the small samples, our omnibus tests for the ANOVAs had a power less than .80. Knowing this, there exists the possibility of committing type II error. Therefore, the differences with other comorbidity measures we described affect our ability to make inferences to the overall population. However, most joint arthroplasty cases described in the literature are similar in severity to our cohort. Another important limitation is that we did not validate the accuracy in the APR-DRGs coding system, which is a multistep processes in our hospital and completed by nonmedical personnel. Moreover, our study captures only the inpatient cost immediately after hospitalization and not after discharge. Also, we used nonprobability sampling techniques; therefore, we recommend much larger investigations using probability methods in a prospective design while addressing our limitations and either refuting or accepting (verifying) our results.

This study is in agreement with Bozic et al [2] who reported that APR-DRGs were predictive of hospital cost in patients receiving lower extremity total joint arthroplasty. In our study, individuals in higher SOI levels and higher ROM levels had higher fixed and variable costs compared to individuals with low SOI. Moreover, those patients who were in the highest SOI and ROM levels had longer length of stay compared to patients in lower levels. Shah and Vail [10] reported that a significant increase in cost corresponded with more severe comorbid illness as measured by APR-DRGs. However, in our analysis, correlations of comorbidity categories with outcome scores were poor. From our data and other published studies, it is clear that the APR-DRG adjustments are best suited in identifying patients who incur a higher hospital cost and patients who present with worse function.

References

1. Bozic KJ, Chiu V. Quality measurement and public reporting in total joint arthroplasty. J Arthroplasty 2008; 23(6 Suppl 1):146.
2. Bozic KJ, Rubash HE, Sculco TP, et al. An analysis of Medicare payment policy for total joint arthroplasty. J Arthroplasty 2008;23(6 Suppl 1):133.
3. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. J Bone Joint Surg Br 1972;54:61.
4. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40:373–83.
5. Mehta RH, Liang L, Karve AM, et al. Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. JAMA 2008; 300(16):1897.
6. Wasielewski RC, Weed H, Prezioso C, et al. Patient comorbidity: relationship to outcomes of total knee arthroplasty. Clin Orthop Relat Res 1998;85.
7. Ethgen O, Bruyere O, Richy F, et al. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. J Bone Joint Surg Am 2004;86-A:963.
8. Cullen DJ, Apolone G, Greenfield S, et al. ASA Physical Status and age predict morbidity after three surgical procedures. Ann Surg 1994;220:3.
9. Greenfield S, Apolone G, McNeil BJ, et al. The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. Med Care 1993;31:141.
10. Shah AN, Vail TP, Taylor D, et al. Comorbid illness affects hospital costs related to hip arthroplasty: quantification of health status and implications for fair reimbursement and surgeon comparisons. J Arthroplasty 2004;19:700.