

## Do Claims-based Comorbidities Adequately Capture Case Mix for Surgical Site Infections?

Hilal Maradit Kremers MD, Laura W. Lewallen MD, Brian D. Lahr MS, Tad M. Mabry MD, James M. Steckelberg MD, Daniel J. Berry MD, Arlen D. Hanssen MD, Elie F. Berbari MD, Douglas R. Osmon MD

Received: 15 July 2014 / Accepted: 20 November 2014 / Published online: 6 December 2014  
© The Association of Bone and Joint Surgeons® 2014

### Abstract

**Background** There is increasing interest in using administrative claims data for surveillance of surgical site infections in THAs and TKAs, but the performance of claims-based models for case-mix adjustment has not been well studied. Performance of claims-based models can be improved with the addition of clinical risk factors for surgical site infections.

**Questions/purposes** We assessed (1) discrimination and calibration of claims-based risk-adjustment models for surgical site infections; and (2) the incremental value of adding clinical risk factors to claims-based risk-adjustment models for surgical site infections.

**Patients and Methods** Our study included all THAs and TKAs performed at a large tertiary care hospital from January 1, 2002 to December 31, 2009 (total n = 20,171 procedures). Revision procedures for infections were excluded. Comorbidity data were ascertained through

administrative records and classified by the Charlson comorbidity index. Clinical details were obtained from the institutional joint registry and patients' electronic health records. Cox proportional hazards regression models were used to estimate the 1-year risk of surgical site infections with a robust sandwich covariance estimator to account for within-subject correlation of individuals with multiple surgeries. The performance of claims-based risk models with and without the inclusion of four clinical risk factors (morbid obesity, prior nonarthroplasties on the same joint, American Society of Anesthesiologists score, operative time) was assessed using measures of discrimination (C statistic, Somers'  $D_{xy}$  rank correlation, and the Nagelkerke  $R^2$  index). Furthermore, calibrations of claims-based risk models with and without clinical factors were assessed graphically by plotting the smoothed trends between model predictions and empirical rates from Kaplan-Meier.

**Results** Discrimination of the claims-based risk models was moderate for the THA (C statistic = 0.662,  $D_{xy}$  = 0.325,  $R^2$  = 0.028) and TKA (C statistic = 0.621,  $D_{xy}$  = 0.241,  $R^2$  = 0.017) cohorts. Inclusion of four clinical risk factors improved discrimination in both cohorts with significant improvement in the C statistic in the THA cohort (C statistic = 0.043; 95% CI, 0.012–0.074) and in the TKA cohort (C statistic = 0.027; 95% CI, 0.007–0.047). Visual inspection suggested that calibration of the claims-based risk models was adequate and comparable to that of models which included the four additional clinical factors.

**Conclusions** Claims-based risk-adjustment models for surgical site infections in THA and TKA appear to be adequately calibrated but lack predictive discrimination, particularly with TKAs. The addition of clinical risk factors improves the discriminative ability of the models to a moderate degree; however, addition of clinical factors did not change calibrations, as the models showed reasonable

---

Each author certifies that he or she, or a member of their immediate family, has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

---

H. Maradit Kremers (✉), L. W. Lewallen,  
B. D. Lahr, T. M. Mabry, J. M. Steckelberg,  
D. J. Berry, A. D. Hanssen, E. F. Berbari, D. R. Osmon  
Mayo Clinic, 200 First Street SW, Harwick 6-69,  
Rochester, MN 55905, USA  
e-mail: Maradit@mayo.edu

degrees of calibration. When used in the clinical setting, the predictive performance of claims-based risk-adjustment models may be improved further with inclusion of additional clinical data elements.

## Introduction

THAs and TKAs are common procedures in the United States [32]. Approximately 0.5% to 3% of patients undergoing THAs and TKAs develop surgical site infections [20, 22, 25, 26]. Variation in surgical site infections rates across hospitals is partly attributable to differences in case mix and partly to differences in prevention practices. Therefore, these infection rates are considered an important hospital quality measure and part of value-based purchasing criteria.

Although there is widespread consensus that public reporting of surgical site infections requires adjustment for case-mix variability in patient populations across hospitals, identifying the optimal approach to risk adjustment remains a difficult process [2, 8]. Traditionally, a generic risk index had been used to adjust for surgical site infection rates as part of the CDC National Healthcare Safety Network (NHSN) tracking system [10]. Procedure-specific versions have been adopted by the CDC more recently and the new versions include additional adjustors to better account for the case mix of patients undergoing THA and TKA across hospitals [24]. However, participation in the NHSN tracking system is not universal. In addition to the revised NHSN risk-adjustment method, the utility of Medicare claims data for automated surveillance of surgical site infections has been investigated [9, 16]. In particular, the Yale New Haven Health Services Corporation/Center for Outcomes Research and Evaluation developed a hospital-level, risk-standardized complication-rate measure for primary THAs and TKAs using 29 hierarchical condition-categories-defined comorbidities for case-mix risk-adjustment [16]. In another study, Bozic et al. [5–7] relied on the Medicare 5% claims data sample to develop claims-data-based risk-prediction models for periprosthetic joint infections. Medicare data-based analyses were based on patient age, sex, and claims-based comorbidities and did not include other significant clinical predictors of surgical site infections, such as surgical history of the joint, the American Society of Anesthesiologists (ASA) score, or immunosuppressed states [3, 4, 12, 14, 19, 25, 34].

The fundamental premise of all the efforts is that administrative claims data are readily available, contain a wealth of information, and can be used for surveillance of surgical site infections using claims data-derived comorbidities as a case-mix adjustor and risk prediction. However, claims data lack basic clinical details, and the

performance of comorbidities in predicting surgical site infection risk may be compromised by the absence of procedure- and patient-specific risk factors not adequately captured in claims data. So far, the performance of claims-based risk-adjustment models for surgical site infections has not been tested against clinically derived risk models.

With this background, our objective was to evaluate the calibration and discrimination of claims-based risk-adjustment models and the incremental value of adding clinical data to the claims-based models. The fundamental question we sought to address was whether claims-based comorbidities adequately capture case-mix for surgical site infections in THAs and TKAs. Specifically, we sought to assess (1) the discrimination and calibration of claims-based risk adjustment models for surgical site infections, and (2) the incremental value of adding clinical risk factors to claims-based risk-adjustment models for surgical site infections.

## Patients and Methods

This retrospective study included all patients who underwent THA and TKA (primary and revision, excluding revision for surgical site infections) at the Mayo Clinic in Rochester Minnesota, between January 1, 2002 and December 31, 2009. Patients who had denied authorization for use of their medical records in research studies, those who underwent hip hemiarthroplasty, partial knee replacement procedures, and revision procedures for prosthetic joint infections were excluded from the study.

Patient comorbidities recorded at the index hospitalization were identified using the ICD-9 diagnoses from index admission and were classified using the Charlson comorbidity index [11]. Because the majority of surgeries were the only encounter for referral patients operated on at our institution, it was not possible to identify comorbidities from previous (12-month) data. Perioperative and intraoperative surgery-specific data were obtained from the institutional total joint registry and electronic health records and included four clinical risk factors previously shown to be predictive of surgical site infections [4]. The four factors were prior surgeries on the same joint, length of operative time, ASA score, and morbid obesity. Electronic health records were used to capture patient height and weight measurements and BMI at the time of surgery. Morbid obesity was defined as having a BMI greater than 40 kg/m<sup>2</sup>.

The primary clinical endpoint for our study was surgical site infections within 1 year after surgery. Potential surgical site infections, including superficial and deep infections, were identified using the total joint registry followup data and the institutional Infection Control Database. Medical records of all patients with potential surgical site infections

were reviewed manually using a pretested data collection form to validate diagnoses and collect required data elements so that infections could be classified according to previously proposed criteria [4, 27, 28].

Descriptive statistics of study population baseline characteristics were reported as mean  $\pm$  SD or count (percentage), as appropriate. Cox proportional hazards regression model was used to test the association of claims-based and clinical-based risk factors with 1-year postsurgical risk of surgical site infections. Because data were expressed at the joint level with some individuals having multiple surgeries, a robust sandwich covariance estimator was used to correct the working independence model for within-subject correlation. Multivariable claims-based risk models were fit separately on the THA and TKA cohorts, using an a priori set of predictor variables that included patient age, sex, type of surgery (primary versus revision surgery), and 16 individual Charlson index comorbidities [5, 6, 9]. We then fit a second set of models, which included the same claims-based predictors and, additionally, four clinical predictors—morbid obesity (BMI  $> 40$  kg/m<sup>2</sup>), prior surgeries on the same joint, ASA score, and length of operative time.

Assessment of overall model performance was based on predictive discrimination (that is, the model's ability to distinguish subjects who had surgical site infections from those who did not) as measured by the concordance (C) statistic, Somers'  $D_{xy}$  rank correlation, and the Nagelkerke  $R^2$  index. The C statistic from a Cox proportional hazards model represents the proportion among all usable pairs of patients in which the patient with longer observed survival was predicted by the model. A pair of patients is considered usable if both subjects have surgical site infections but at different times, or if one has a surgical site infection and the other is censored at an equal or later followup time. A C statistic of 0.5 corresponds to a noninformative model, whereas a value of 1.0 corresponds to a perfectly discriminating model. The  $R^2$  index is another unitless measure of the model's predictive ability, where higher values indicate more variation in the survival outcome is explained by the model. For internal validation of the models, we used 200 bootstrap resamples (each selected via random sampling with replacement from the original cohort of an equal sample size) to estimate the bias attributable to overfitting or overoptimism in each model and then recalculated the measures of model performance correcting for this bias. For formal comparison of the performance of the claims-based model with and without inclusion of the clinical predictors, we used an integrative discrimination improvement to show how far, on average, individuals were moving (in an appropriate direction) along the continuum of predicted risk. The improved discrimination, as a result of adding the four clinical predictors to the model, was estimated separately in groups

of patients with and without surgical site infections. Higher integrative discrimination improvement among patients who had surgical site infections indicated improvement in sensitivity, whereas a higher integrative discrimination improvement among patients without surgical site infections indicated improvement in specificity [23]. These then were combined to derive an aggregate estimate of integrative discrimination improvement based on the equally weighted sum of improvement, which was tested for significance using a one sample t-test.

To assess calibration—the model's ability to predict accurately the subsequently observed surgical site infection risk—the model survival predictions were binned into deciles and plotted against empiric rates free of surgical site infections, estimated using the Kaplan-Meier method. A calibration curve based on a nonparametric, locally estimated scatterplot smoothing estimator was used to show how closely the predicted survival values related to the observed frequency without surgical site infections. A curve that deviated appreciably from the 45° line of identity (ideal calibration) would show evidence that the model was miscalibrated. Data points to the right of the identity line indicate overestimation of surgical site infection survival by the model (that is, the model would underestimate the risk of surgical site infections developing), whereas points to the left indicate underestimation. All analyses were performed using SAS<sup>®</sup> software, version 9.3 (SAS Institute, Cary, NC, USA).

Our study included 9720 THAs (2515 revisions) in 8270 patients and 10,451 TKAs (1522 revisions) in 8253 patients (Table 1). Overall, mean patient age ( $\pm$  SD) at the time of surgery was 66.2 ( $\pm$  12.6) years and patients who underwent TKAs were slightly older. Less than 1/2 (45%) of the patients were men. Mean operative time was longer for revision procedures and ASA scores were slightly higher at the time of revisions compared with primary procedures. Mean BMI was slightly higher in patients who underwent TKAs. The overall proportion of morbidly obese patients (BMI  $> 40$  kg/m<sup>2</sup>) was 9% and there were more morbidly obese patients (12%) in the primary TKA group. A total of 368 surgical site infections occurred during the 1-year window after surgery. Of these, 192 were prosthetic joint infections, as classified by the Infectious Diseases Society of America criteria [27]. The 1-year cumulative incidences of all surgical site infections were 1.31%, 2.83%, 1.88%, and 2.98% for primary THA, revision THA, primary TKA, and revision TKA, respectively.

## Results

The adjusted hazard ratios for risk of surgical site infections were calculated for the claims-based and clinical

**Table 1.** Baseline characteristics of study cohorts

Variable	Primary THA (n = 7205)	Revision THA (n = 2515)	Primary TKA (n = 8929)	Revision TKA (n = 1522)	Total (n = 20,171)
Age (mean years ± SD)	64.0 ± 14.1	66.5 ± 13.9	67.8 ± 10.5	67.1 ± 11.8	66.2 ± 12.6
Men, n (%)	3496 (49)	1117 (44)	3899 (44)	665 (44)	9177 (45)
Prior operation, n (%)	689 (10)	2514 (100)	2295 (26)	1522 (100)	7020 (35)
Operative time (mean minutes ± SD)	132.9 ± 56.3	194.2 ± 89.6	143.4 ± 60.7	182.6 ± 79.3	148.9 ± 68.5
ASA score, n (%)					
1	353 (5)	68 (3)	161 (2)	29 (2)	611 (3)
2	4342 (61)	1311 (53)	5232 (60)	862 (58)	11747 (59)
3	2347 (33)	1074 (43)	3301 (38)	584 (39)	7306 (37)
4	60 (1)	31 (1)	50 (1)	12 (1)	153 (1)
BMI* (mean kg/m <sup>2</sup> ± SD)	29.7 ± 6.3	29.0 ± 6.0	32.2 ± 6.7	31.8 ± 6.3	30.9 ± 6.6
Morbid obesity*, n (%)	476 (7)	123 (5)	1069 (12)	158 (10)	1826 (9)
Comorbidities <sup>†</sup> , n (%)					
Myocardial infarction	539 (7)	212 (8)	821 (9)	97 (6)	1669 (8)
Congestive heart failure	669 (9)	243 (10)	988 (11)	124 (8)	2024 (10)
Peripheral vascular disease	717 (10)	200 (8)	1011 (11)	100 (7)	2028 (10)
Cerebrovascular disease	928 (13)	268 (11)	1368 (15)	158 (10)	2722 (13)
Dementia	69 (1)	25 (1)	77 (1)	11 (1)	182 (1)
Pulmonary disease	1059 (15)	315 (13)	1695 (19)	217 (14)	3286 (16)
Ulcers	636 (9)	187 (7)	1092 (12)	135 (9)	2050 (10)
Mild liver disease	230 (3)	85 (3)	446 (5)	41 (3)	802 (4)
Severe liver disease	53 (1)	23 (1)	137 (2)	8 (1)	221 (1)
Diabetes mellitus	1087 (15)	320 (13)	1829 (20)	271 (18)	3507 (17)
Diabetes with organ damage	300 (4)	73 (3)	564 (6)	55 (4)	992 (5)
Hemiplegia	68 (1)	21 (1)	98 (1)	7 (0)	194 (1)
Renal disease	879 (12)	274 (11)	1267 (14)	149 (10)	2569 (13)
Metastatic tumors	407 (6)	96 (4)	559 (6)	59 (4)	1121 (6)
Rheumatologic disease	564 (8)	266 (11)	797 (9)	136 (9)	1763 (9)
Cancer	1613 (22)	389 (15)	2331 (26)	248 (16)	4581 (23)

\* BMI defined using actual patient height and weight measurements at time of surgery. Morbid obesity = BMI > 40 kg/m<sup>2</sup>; <sup>†</sup>Comorbidities defined per the Charlson comorbidity index; ASA = American Society of Anesthesiologists.

multivariable models, separately for THAs and TKAs. Of the 16 comorbidities included in the multivariable models for patients who underwent THAs (Table 2), pulmonary disease, mild liver disease, diabetes mellitus, and rheumatologic diseases were significantly associated with surgical site infections, once adjustments were made for patient age, sex, and revision status (the claims model). Of the associated comorbidities, mild liver and rheumatologic diseases were associated with surgical site infections when further adjustments for the four clinical factors (the clinical model) were performed. The four clinical factors contributed to the model, either in terms of individual effects (all except operative time), and when tested as a combined effect based on model-log likelihood (37.4;  $p < 0.001$ ). In the TKA cohort (Table 3), renal disease was the only comorbidity significantly associated with surgical site infections in the claims and clinical models. In the clinical

model, morbid obesity was the only clinical predictor of surgical site infections, although the overall contribution from all four clinical factors was shown to be significant based on log likelihood (17.0;  $p = 0.009$ ).

As indicated by the performance statistics, the models showed better performance in the THA cohort than the TKA cohort (Table 4). The original fit of the claims-based THA model, which included patient age, sex, surgery type (primary versus revision), and Charlson comorbidities, had a C statistic of 0.662 (although an overfitting-corrected estimate of 0.629 may provide a more accurate measure of the model discrimination in our cohort). The C statistic increased to 0.706 (bias-corrected, 0.665) with the addition of the four clinical risk factors (difference, 0.043; 95% CI, 0.012–0.074;  $p = 0.006$ ). The C statistic of 0.621 (bias-corrected, 0.585) from the claims-based TKA model improved significantly to 0.648 (bias-corrected, 0.606) with the addition of

**Table 2.** Hazard ratios for 1-year risk of surgical site infections in THA

Factor	Hazard ratio (95% CI)**	
	Claims model	Claims + clinical model
Age (per 10 years increase)	0.91 (0.82–1.02)	0.89 (0.79–1.01)
Male gender	0.89 (0.64–1.23)	0.86 (0.62–1.19)
Prior hip replacement	2.25 (1.65–3.08)*	1.14 (0.65–2.00)
Myocardial infarction	1.01 (0.55–1.86)	0.99 (0.55–1.80)
Congestive heart failure	1.13 (0.62–2.05)	1.07 (0.61–1.89)
Peripheral vascular disease	0.70 (0.38–1.27)	0.71 (0.39–1.28)
Cerebrovascular disease	1.06 (0.66–1.72)	1.01 (0.63–1.63)
Dementia	1.28 (0.31–5.24)	1.22 (0.30–5.04)
Pulmonary disease	1.56 (1.03–2.36)*	1.43 (0.95–2.14)
Ulcers	0.91 (0.54–1.55)	0.89 (0.53–1.51)
Mild liver disease	2.78 (1.62–4.75) *	2.41 (1.41–4.13)*
Severe liver disease	0.59 (0.14–2.45)	0.68 (0.17–2.73)
Diabetes Mellitus	1.56 (1.02–2.38)*	1.22 (0.79–1.89)
Diabetes with organ damage	1.17 (0.57–2.39)	1.22 (0.60–2.50)
Hemiplegia	1.15 (0.26–5.07)	1.11 (0.25–4.97)
Renal Disease	1.21 (0.76–1.90)	1.13 (0.73–1.77)
Metastatic tumors	1.33 (0.70–2.55)	1.26 (0.67–2.38)
Rheumatologic disease	2.01 (1.32–3.05)*	1.89 (1.24–2.88)*
Cancer	1.18 (0.79–1.76)	1.20 (0.80–1.79)
Morbid obesity		2.09 (1.29–3.38)*
Prior operation		1.99 (1.15–3.46)*
ASA		
1		1.0 (reference)*
2		2.46 (0.61–9.93)
3		4.80 (1.18–19.61)
4		1.92 (0.17–21.64)
Operative time (per 10 minutes)		1.01 (0.98–1.03)

\* Significant at an alpha level of 0.05; \*\*estimates derived from multivariable Cox regression models with and without clinical risk factors; ASA = American Society of Anesthesiologists.

the four clinical risk factors (difference, 0.027; 95% CI, 0.007–0.047;  $p = 0.009$ ). Furthermore, adding the four clinical factors to the claims-based model translated to an improvement in the integrative discrimination improvement among patients who had THA (integrative discrimination improvement, 0.37%; 95% CI, 0.12–0.62;  $p = 0.003$ ), particularly for the subgroup with surgical site infections (that is, improvement in sensitivity) (Fig. 1). In contrast, the integrative discrimination improvement was not significantly improved in the models for TKA (Fig. 2), neither in the overall patient group nor in the subset of patients with surgical site infections.

**Table 3.** Hazard ratios for 1-year risk of surgical site infections in TKA

Factor	Hazard ratio (95% CI)	
	Claims model	Claims + clinical model
Age (per 10 years increase)	0.76 (0.67–0.86)*	0.77 (0.66–0.89)*
Male gender	1.14 (0.86–1.51)	1.17 (0.88–1.56)
Prior knee replacement	1.64 (1.17–2.30)*	1.57 (1.02–2.42)*
Myocardial infarction	1.14 (0.69–1.89)	1.08 (0.66–1.78)
Congestive heart failure	1.61 (1.00–2.58)	1.49 (0.93–2.38)
Peripheral vascular disease	0.93 (0.58–1.47)	0.93 (0.59–1.48)
Cerebrovascular disease	1.02 (0.69–1.52)	1.00 (0.67–.49)
Dementia	–	–
Pulmonary disease	0.96 (0.67–1.38)	0.90 (0.62–1.30)
Ulcers	1.07 (0.70–1.65)	1.06 (0.70–1.63)
Mild liver disease	1.61 (0.93–2.79)	1.55 (0.90–2.69)
Severe liver disease	1.55 (0.66–3.63)	1.51 (0.63–3.60)
Diabetes mellitus	0.86 (0.57–1.30)	0.73 (0.48–1.11)
Diabetes with organ damage	0.84 (0.43–1.64)	0.81 (0.41–1.60)
Hemiplegia	1.64 (0.62–4.37)	1.68 (0.63–4.50)
Renal Disease	1.72 (1.14–2.62)*	1.67 (1.10–2.55)*
Metastatic tumors	1.37 (0.78–2.42)	1.37 (0.78–2.42)
Rheumatologic disease	0.98 (0.61–1.57)	1.00 (0.62–1.60)
Cancer	1.10 (0.77–1.57)	1.10 (0.77–1.58)
Morbid obesity		1.78 (1.21–2.61)*
Prior operation		1.06 (0.73–1.52)
ASA		
1		1.0 (reference)
2		1.00 (0.36–2.82)
3		1.39 (0.47–4.11)
4		2.27 (0.46–11.15)
Operative time (per 10 minutes)		1.00 (0.98–1.02)

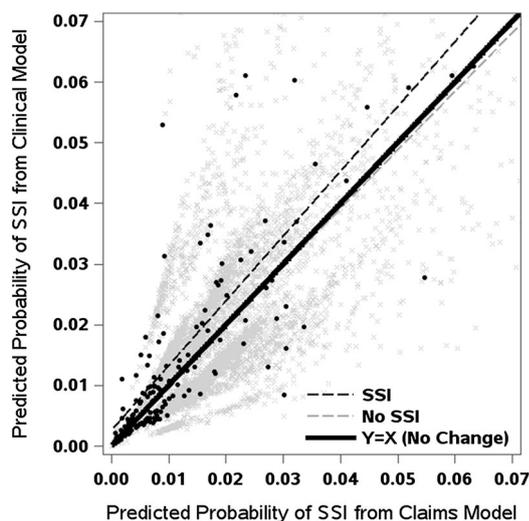
\* Significant at an alpha level of 0.05; \*\*estimates derived from multivariable Cox regression models with and without clinical risk factors; ASA = American Society of Anesthesiologists.

The claims-based models appeared reasonably calibrated and the addition of clinical factors did not result in noticeable improvement in calibration in either THA (Fig. 3) or TKA (Fig. 4). Visual inspection of the empirical free-of-surgical site infections rates by prediction deciles and the calibration curve from a nonparametric locally estimated scatterplot smoothing estimator suggested no clear signs of miscalibration in either the claims-based or the clinical models. The calibration curves of the two models looked fairly similar. The models appeared to be reasonably calibrated without noticeable improvement once the clinical factors were added to the models.

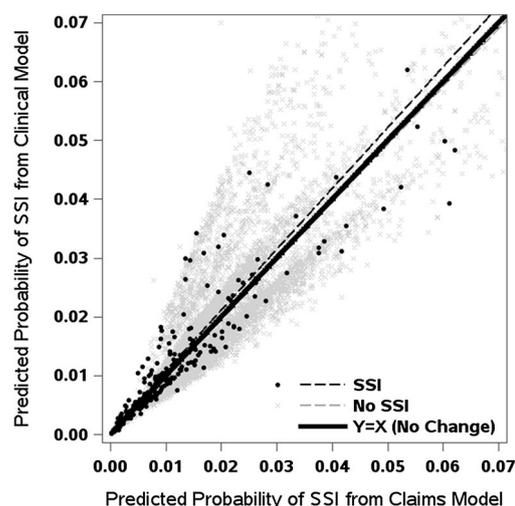
**Table 4.** Performance of claims-based risk-prediction models

Performance statistic	THA		TKA	
	Claims model	Claims + clinical model	Claims model	Claims + clinical model
Model accuracy index (original/bias-corrected)				
C statistic	0.662/0.629	0.706/0.665	0.621/0.585	0.648/0.606
Somers' $D_{xy}$ statistic	0.325/0.258	0.411/0.330	0.241/0.169	0.295/0.212
$R^2$ (%)	0.028/0.013	0.042/0.021	0.017/0.007	0.023/0.009
Comparison of models (improvement, 95% CI or p value)				
Difference in C statistic	0.043 (0.012–0.074)*		0.027 (0.007–0.047)*	
Difference in log-likelihood	37.4 (p < 0.001)		17.0 (p = 0.009)	
Integrated discrimination improvement in subjects with surgical site infection	0.36% (0.11%–0.61%)		0.09% (–0.02% to 0.20%)	
Integrated discrimination improvement subjects without surgical site infection	0.01% (–0.01% to 0.03%)		0.00% (–0.01% to 0.02%)	
Aggregate integrated discrimination improvement (equal-weighting of two groups)	0.37% (0.12%–0.62%)		0.09% (–0.02% to 0.21%)	

\* CIs calculated using the jackknife.



**Fig. 1** For patients who had THAs, improvement with the addition of clinical risk factors is reflected by predicted values and a regression line above the 45° line of identity (ie, no change) in patients with surgical site infections (SSI) or below this line in patients without surgical site infections.



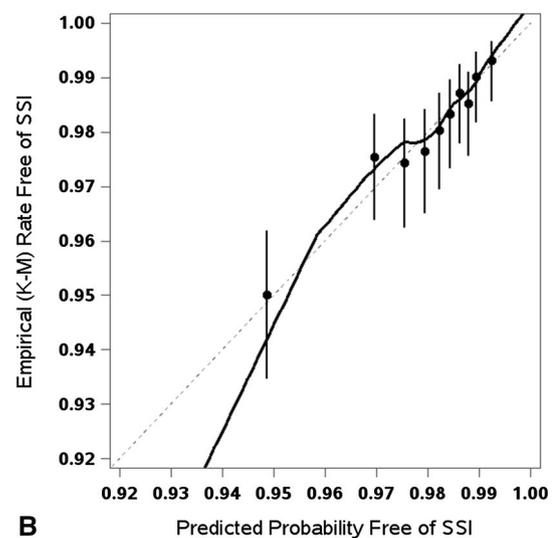
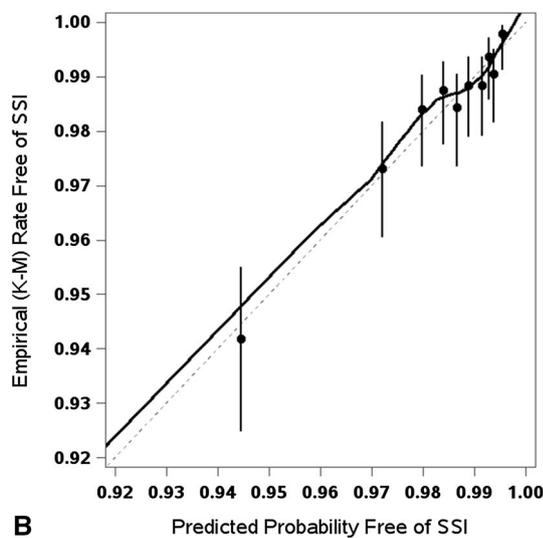
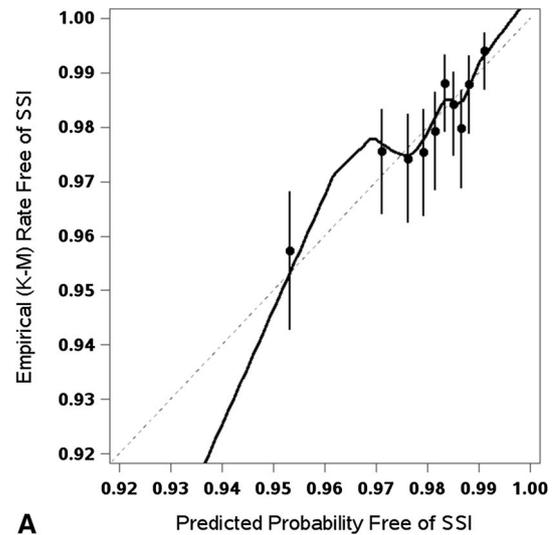
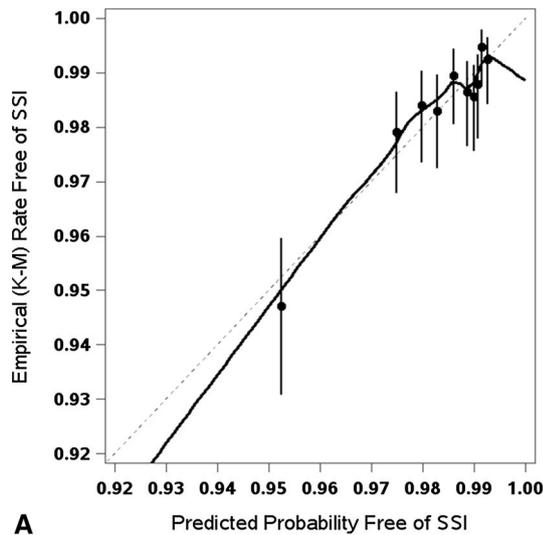
**Fig. 2** For patients who had TKAs, improvement with the addition of clinical risk factors is reflected by predicted values and a regression line above the 45° line of identity (ie, no change) in those with surgical site infections (SSI) or below this line in those without surgical site infections.

## Discussion

There is increasing interest in using administrative-claims data for surveillance of surgical site infections but the performance of claims-based models for case-mix adjustment has not been well studied. In a large cohort of patients who underwent total joint replacement at one tertiary care institution, we examined the predictive performance of claims data-based comorbidities as a case-mix adjustor for

surgical site infections and the incremental value of adding clinical risk factors. Our findings suggest that the claims-based risk-adjustment models for surgical site infections are well calibrated but lack predictive discrimination. Discrimination can be improved by the addition of clinical risk factors.

Our study results should be interpreted in light of some potential limitations. First, our study was confined to procedures performed at one hospital, which can restrict power



**Fig. 3A–B** The calibration plots show the observed versus predicted probability of surgical site infections (SSI) for THAs using the (A) claims-based model and (B) the claims + clinical risk factors model. The dots represent Kaplan-Meier (K-M) survival estimates with the vertical bars showing the respective 95% CIs.

**Fig. 4A–B** The calibration plots show the observed versus predicted probability of surgical site infections (SSI) for TKAs using the (A) claims-based model and (B) the claims + clinical risk factors model. The dots represent Kaplan-Meier (K-M) survival estimates with the vertical bars showing the respective 95% CIs.

and generalizability of findings to other institutions. However, this is also a strength of the study because confounding by unknown and known healthcare delivery factors, such as infection-prevention practices, is minimized. Ideally, future multicenter studies will take into account hospital-level healthcare delivery factors when assessing the performance of risk prediction models. Second, our method for ascertainment of comorbidities and surgical site infections outcomes is different than those of previous analyses. Similar to other investigators [5–7, 9, 16], we relied on electronically available administrative data to ascertain comorbidities but our time window was limited to the index hospitalization. This is in contrast to a

Medicare study which ascertained comorbidities from the 1-year period before admission [16]. Therefore, we may have under-ascertained some comorbidities. Despite this, a qualitative comparison of the prevalence of comorbidities revealed minimal differences. In addition, the types of comorbidities we included are different from other claims-based analyses [5–7, 16]. For example, the Centers for Medicare & Medicaid Services complication measure [16] grouped comorbidities using 29 condition categories, whereas the analyses by Bozic et al. [5–7] were based on the Elixhauser method. Other important differences in our study include the exclusion criteria, outcome definitions, and the time for ascertainment of surgical site infections.

Despite these differences in methods, unique strengths of our study are complete ascertainment and chart review-based validation of all surgical site infection outcomes, which minimized measurement bias. Owing to active followup of patients through the institutional total joint registry, we were able to identify all surgical site infections during the 1 year after surgery. In addition, partly because of the availability of unique institutional resources, we were able to ascertain and examine difficult to obtain clinical risk factors known to be associated with surgical site infections. This rarely is possible in other settings. Finally, although our cohorts are large, it is questionable whether the number of surgical site infections is adequate to ensure a reliable model of this many candidate predictor variables and avoid overfitting. We assessed this using internal validation and found a modest amount of overfitting based on model C statistics which were overoptimistic by approximately 0.03 to 0.04. Since overfitting tends to produce overestimated effects, we estimated the calibration slope in the bootstrap procedure and found that model accuracy would benefit from some shrinkage in these regression coefficients (by approximately 20% to 25%) which further suggests overspecification in the modeling.

Our results indicate that the discrimination (ability to separate patients who do and do not experience surgical site infections) of claims-based models is poor (bias-corrected C statistics, 0.629 for THA and 0.585 for TKA), but substantially improved with the addition of four well-established clinical risk factors for surgical site infections, particularly in patients who underwent THA. Nevertheless, discrimination of the clinical models remained modest (bias-corrected C statistics were less than 0.7), suggesting that other clinical risk factors could be incorporated to risk-adjustment models of surgical site infections to further improve prediction.

Surgical site infections, particularly, prosthetic joint infections, cause substantial patient morbidity and mortality, and contribute to arthroplasty-related healthcare costs [21, 35]. Although the overall incidence of surgical site infections is low, the absolute number is expected to increase in the future owing to increased volume of total joint replacement procedures and growth in the proportion of high-risk patients. Numerous surgical site infections potentially are preventable through effective management of patient risk factors and implementation of hospital-based infection-prevention practices. Surveillance of surgical site infections rates is considered an important quality measure for hospital profiling and future value-based purchasing and pay-for-performance programs. However, not all patients have the same baseline risk for surgical site infections and hospitals and surgeons differ in the case mix of their patients. Therefore, risk adjustment for case-mix differences is an essential component of comparisons of

surgical site infections across hospitals. Apart from issues related to case-mix adjustment, risk prediction of surgical site infections for an individual patient can define specific prevention strategies. Reliance on claims data is becoming increasingly popular because the data are readily available and offer the ability to conduct surveillance activities with minimal additional data-collection requirements. However, many of the strong predictors of surgical site infections in total joint replacement are intrinsic patient factors and/or surgical factors that are not adequately captured in claims data. Coupled with the inadequacy of claims data in terms of completeness and coding errors, there are concerns regarding the potential consequences of using claims data for case-mix adjustment in surveillance of surgical site infections in total joint replacement [15, 29]. Therefore, it is important that the findings of claims-based analyses are validated against high-quality, clinically derived data, because prior studies of various medical conditions and cardiac procedures indicate that the concordance between administrative claims data and medical-record documentation may vary substantially, and adding clinical data to claims data can significantly improve risk prediction and also alter performance ratings of hospitals [13, 17, 30].

Our analyses are an external assessment of claims-based risk models for surgical site infections in one hospital, focusing on two important measures of predictive performance [33]. Discrimination is the ability of a model, using a set of predictor variables, to accurately separate patients who did and did not experience surgical site infections. The C statistic is the standard approach to quantifying discrimination, where a value of 0.5 indicates that discrimination is no better than chance, whereas larger values indicate better discrimination. We found the discrimination of claims-based models (especially in TKA) to be lacking, with a bias-corrected C statistic of approximately 0.6. This may be attributable to two factors: lack of prediction of the claims-based risk models and/or a more homogeneous case mix in our cohort (that is, low variability of known and unknown risk factors for surgical site infections and consequently, patients being more alike). Only a few of the claims-based comorbidities in our study were found to be significantly associated with surgical site infections, indicating lack of model prediction. Surprisingly, three of four clinically well-recognized risk factors for surgical site infections [1] were not found to be predictive for surgical site infections among patients who underwent TKA. Furthermore, our study cohort possibly had a more severe case mix and was more homogeneous than other large multicenter cohorts—in terms of risk factors and rate of surgical site infections—such as the Medicare population. Nevertheless, it is not fair to discard claims-based models based on the low C statistic in one external dataset, plus the C statistic is only one of many

discrimination measures. The integrative discrimination improvement analysis indicates substantial improvement in risk classification among individuals with surgical site infections in the THA cohort. Although the average integrative discrimination improvement of 0.36% in patients with surgical site infections seems like a negligible amount, one must consider this is on a scale of absolute risk and generally the overall risk of surgical site infections is low (approximately 1.7%). Additional studies are warranted to document how comorbidities and clinical risk-factor profiles vary across hospitals and the implications for hospital performance profiling, if any.

To our knowledge, no previous claims-based analyses reported on the performance of the prediction models used, therefore, it is not possible to compare our findings with previously published analyses. Furthermore, study definitions of comorbidities are not alike, except in one study [9]. Although the prevalence of some more common comorbidities and the risk estimates appear to be comparable, there are important differences. For example, the relatively low prevalence of dementia and congestive heart failure may be a reflection of the younger patients and higher proportion of women in our cohort compared with the Medicare datasets. Furthermore, residual confounding is a potential limitation with ICD-based comorbidity definitions owing to coding inaccuracy, completeness, and absence of severity information [18]. An important unexpected finding of our study is that the risk factors for surgical site infections and performance of prediction and risk-adjustment models may differ depending on type of joint replacement surgery. This was suggested previously [31] and deserves further investigation. If true, analyses of outcome data for surgical site infections from combined cohorts should be stratified by surgery type.

In terms of calibration (agreement between observed and predicted outcomes of surgical site infections), it is encouraging that calibration of claims-based models appeared adequate in THA and TKA cohorts, indicating that the predicted probabilities of surgical site infections from the claims-based models were fairly agreeable with the observed rates of surgical site infections. The addition of the four clinical risk factors provided little incremental improvement. However the relatively low incidence of surgical site infections combined with a lack of predictive discrimination resulted in a rather narrow and skewed spread of the 10-decile risk groups from these models. A model that is well calibrated across the entire range of predictions is particularly important in hospital profiling. For example, if the calibration of a surgical site infections model is poor in the most at-risk patients, then a hospital with a worse case mix can be unfairly labeled a poor performer. Our study is limited to one institution and further validation studies in larger, multicenter datasets are

needed to better scrutinize the implications of calibration for hospital profiling. Further validation studies are warranted to determine how well claims-based and clinical models perform as a case-mix adjustor across hospitals. The addition of clinical risk factors to claims models can significantly improve discrimination, which suggests that if used in the clinical setting, a higher degree of accuracy in predicting surgical site infections risk is achievable.

Claims-based risk-adjustment models for surgical site infections appeared to be well calibrated in our large cohort of patients who underwent total joint replacements; however, the models lacked strong predictive discrimination. Our findings underscore the importance of external validation studies to better understand the validity and generalizability of claims-based risk models for surveillance of surgical site infections in total joint replacements, and how they can be interpreted when assessing case-mix adjustment and risk prediction in the clinic.

## References

1. Aggarwal VK, Tischler EH for Workgroup 1 (Leaders: Lautenbach C, Williams GR Jr; Delegates: Abboud JA, Altena M, Bradbury T, Calhoun J, Douglas D, Del Gaizo DJ, Font-Vizcarra L, Huotari K, Kates S, Koo KH, Mabry TM, Moucha CS, Palacio JC, Peel TN, Poolman RW, Robb WJ 3<sup>rd</sup>, Salvagno R, Seyler T, Skaliczki G, Vasarhelyi EM, Watters WC 3<sup>rd</sup>). Mitigation and Education. In: Parvizi J, Gehrke T, eds. Proceedings of the International Consensus Meeting on Periprosthetic Joint Infection. Available at: <http://www.msiss-na.org/wp-content/themes/msis-temp/pdf/ism-periprosthetic-joint-information.pdf>. Accessed November 14, 2014.
2. Anderson DJ, Chen LF, Sexton DJ, Kaye KS. Complex surgical site infections and the devilish details of risk adjustment: important implications for public reporting. *Infect Control Hosp Epidemiol.* 2008;29:941–946.
3. Berbari EF, Hanssen AD, Duffy MC, Steckelberg JM, Ilstrup DM, Harmsen WS, Osmon DR. Risk factors for prosthetic joint infection: case-control study. *Clin Infect Dis.* 1998;27:1247–1254.
4. Berbari EF, Osmon DR, Lahr B, Eckel-Passow JE, Tsaras G, Hanssen AD, Mabry T, Steckelberg J, Thompson R. The Mayo prosthetic joint infection risk score: implication for surgical site infection reporting and risk stratification. *Infect Control Hosp Epidemiol.* 2012;33:774–781.
5. Bozic KJ, Lau E, Kurtz S, Ong K, Berry DJ. Patient-related risk factors for postoperative mortality and periprosthetic joint infection in medicare patients undergoing TKA. *Clin Orthop Relat Res.* 2012;470:130–137.
6. Bozic KJ, Lau E, Kurtz S, Ong K, Rubash H, Vail TP, Berry DJ. Patient-related risk factors for periprosthetic joint infection and postoperative mortality following total hip arthroplasty in Medicare patients. *J Bone Joint Surg Am.* 2012;94:794–800.
7. Bozic KJ, Ong K, Lau E, Berry DJ, Vail TP, Kurtz SM, Rubash HE. Estimating risk in Medicare patients with THA: an electronic risk calculator for periprosthetic joint infection and mortality. *Clin Orthop Relat Res.* 2013;471:574–583.
8. Brandt C, Hansen S, Sohr D, Daschner F, Ruden H, Gastmeier P. Finding a method for optimizing risk adjustment when comparing surgical-site infection rates. *Infect Control Hosp Epidemiol.* 2004;25:313–318.

9. Calderwood MS, Kleinman K, Bratzler DW, Ma A, Bruce CB, Kaganov RE, Canning C, Platt R, Huang SS, Centers for Disease Control and Prevention Epicenters Program, Oklahoma Foundation for Medical Quality. Use of Medicare claims to identify US hospitals with a high rate of surgical site infection after hip arthroplasty. *Infect Control Hosp Epidemiol.* 2013;34:31–39.
10. CDC Centers for Disease Control and Prevention. National Healthcare Safety Network (NHSN). Available at: <http://www.cdc.gov/nhsn/>. Accessed November 11, 2014.
11. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis.* 1987;40:373–383.
12. Dowsey MM, Choong PF. Obesity is a major risk factor for prosthetic infection after primary hip arthroplasty. *Clin Orthop Relat Res.* 2008;466:153–158.
13. Fonarow GC, Pan WQ, Saver JL, Smith EE, Reeves MJ, Broderick JP, Kleindorfer DO, Sacco RL, Olson DM, Hernandez AF, Peterson ED, Schwamm LH. Comparison of 30-day mortality models for profiling hospital performance in acute ischemic stroke with vs without adjustment for stroke severity. *JAMA.* 2012;308:257–264.
14. Gibbons C, Bruce J, Carpenter J, Wilson AP, Wilson J, Pearson A, Lamping DL, Krukowski ZH, Reeves BC. Identification of risk factors by systematic review and development of risk-adjusted models for surgical site infection Introduction to the research. *Health Technol Assess.* 2011;15:1–156, iii–iv.
15. Gray P, Streed S, Dolan S, Khoury R, Kulich P, Olmsted R. APIC Position Paper: The Use of Administrative (Coding/Billing) Data for Identification of Healthcare-Associated Infections (HAIs) in US Hospitals. 2010. Available at: [http://www.apic.org/Resource\\_/TinyMceFileManager/Advocacy-PDFs/ID\\_of\\_HAIs\\_US\\_Hospitals\\_1010.pdf](http://www.apic.org/Resource_/TinyMceFileManager/Advocacy-PDFs/ID_of_HAIs_US_Hospitals_1010.pdf). Accessed November 11, 2014.
16. Grosso LM, Curtis JP, Lin Z, Geary LL, Vellanky S, Oladele C, Ott LS, Parzynski C, Suter LG, Berhneim SM, Drye EE, Krumholz HM (Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHC/CORE)). Hospital-level Risk-Standardized Complication Rate Following Elective Primary Total Hip Arthroplasty (THA) And/Or Total Knee Arthroplasty (TKA): Measure Methodology Report. Prepared for the Centers for Medicare & Medicaid Services. 2012. Available at: <https://staging.qualitynet.org/>. Accessed November 11, 2014.
17. Hammill BG, Curtis LH, Fonarow GC, Heidenreich PA, Yancy CW, Peterson ED, Hernandez AF. Incremental value of clinical data beyond claims data in predicting 30-day outcomes after heart failure hospitalization. *Circ Cardiovasc Qual Outcomes.* 2011; 4:60–67.
18. Jackson ML, Nelson JC, Jackson LA. Why do covariates defined by International Classification of Diseases codes fail to remove confounding in pharmacoepidemiologic studies among seniors? *Pharmacoepidemiol Drug Saf.* 2011;20:858–865.
19. Jansen E, Huhtala H, Puolakka T, Moilanen T. Risk factors for infection after knee arthroplasty: a register-based analysis of 43,149 cases. *J Bone Joint Surg Am.* 2009;91:38–47.
20. Kurtz SM, Lau E, Schmier J, Ong KL, Zhao K, Parvizi J. Infection burden for hip and knee arthroplasty in the United States. *J Arthroplasty.* 2008;23:984–991.
21. Kurtz SM, Lau E, Watson H, Schmier JK, Parvizi J. Economic burden of periprosthetic joint infection in the United States. *J Arthroplasty.* 2012;27:61–65 e61.
22. Kurtz SM, Ong KL, Lau E, Bozic KJ, Berry D, Parvizi J. Prosthetic joint infection risk after TKA in the Medicare population. *Clin Orthop Relat Res.* 2010;468:52–56.
23. Leening MJ, Cook NR. Net reclassification improvement: a link between statistics and clinical practice. *Eur J Epidemiol.* 2013;28:21–23.
24. Mu Y, Edwards JR, Horan TC, Berrios-Torres SI, Fridkin SK. Improving risk-adjusted measures of surgical site infection for the national healthcare safety network. *Infect Control Hosp Epidemiol.* 2011;32:970–986.
25. Namba RS, Inacio MC, Paxton EW. Risk factors associated with surgical site infection in 30,491 primary total hip replacements. *J Bone Joint Surg Br.* 2012;94:1330–1338.
26. Ong KL, Kurtz SM, Lau E, Bozic KJ, Berry DJ, Parvizi J. Prosthetic joint infection risk after total hip arthroplasty in the Medicare population. *J Arthroplasty.* 2009;24(6 suppl):105–109.
27. Osmon DR, Berbari EF, Berendt AR, Lew D, Zimmerli W, Steckelberg JM, Rao N, Hanssen A, Wilson WR. Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America. *Clin Infect Dis.* 2013;56:e1–e25.
28. Parvizi J, Zmistowski B, Berbari EF, Bauer TW, Springer BD, Della Valle CJ, Garvin KL, Mont MA, Wongworawat MD, Zalavras CG. New definition for periprosthetic joint infection: from the Workgroup of the Musculoskeletal Infection Society. *Clin Orthop Relat Res.* 2011;469:2992–2994.
29. Sarrazin MS, Rosenthal GE. Finding pure and simple truths with administrative data. *JAMA.* 2012;307:1433–1435.
30. Shahian DM, He X, Jacobs JP, Rankin JS, Peterson ED, Welke KF, Filardo G, Shewan CM, O'Brien SM. Issues in quality measurement: target population, risk adjustment, and ratings. *Ann Thorac Surg.* 2013;96:718–726.
31. Song KH, Kim ES, Kim YK, Jin HY, Jeong SY, Kwak YG, Cho YK, Sung J, Lee YS, Oh HB, Kim TK, Koo KH, Kim EC, Kim JM, Choi TY, Kim HY, Choi HJ, Kim HB. Differences in the risk factors for surgical site infection between total hip arthroplasty and total knee arthroplasty in the Korean Nosocomial Infections Surveillance System (KONIS). *Infect Cont Hosp Ep.* 2012;33:1086–1093.
32. Steiner C, Andrews R, Barrett M, Weiss A. *HCUP Projections: Mobility/Orthopedic Procedures 2003 to 2012.* HCUP Projections Report # 2012-03. Online September 20, 2012. U.S. Agency for Healthcare Research and Quality. Available: <http://www.hcup-us.ahrq.gov/reports/projections/2012-03.pdf>. Accessed November 11, 2014.
33. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiol.* 2010;21:128–138.
34. Urquhart DM, Hanna FS, Brennan SL, Wluka AE, Leder K, Cameron PA, Graves SE, Cicuttini FM. Incidence and risk factors for deep surgical site infection after primary total hip arthroplasty: a systematic review. *J Arthroplasty.* 2010;25:1216–1222.
35. Zimlichman E, Henderson D, Tamir O, Franz C, Song P, Yamin CK, Keohane C, Denham CR, Bates DW. Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system. *JAMA Intern Med.* 2013;173:2039–2046.