Spine

SPINE Volume 42, Number 17, pp 1311–1315

HEALTH SERVICES RESEARCH

Predictors of Hospital Readmission and Surgical Site Infection in the United States, Denmark, and Japan

Is Risk Stratification a Universal Language?

Steven Glassman, MD,[*] Leah Y. Carreon, MD, MSc,[*] Mikkel Andersen, MD,[†] Anthony Asher, MD,[‡] Soren Eiskjær, MD,[§] Martin Gehrchen, MD,[¶] Shiro Imagama, MD,[||] Ken Ishii, MD,[**] Takahashi Kaito, MD,[††] Yukihiro Matsuyama, MD,[‡‡] Hiroshi Moridaira, MD,[§§] Praveen Mummaneni, MD,[¶¶] Christopher Shaffrey, MD,[||||] and Morio Matsumoto, MD[**]

Study Design. Retrospective review of three spine surgery databases.

Objectives. The purpose of the present study is to determine whether predictors of hospital readmission and surgical site infection (SSI) after lumbar fusion will be the same in United States, Denmark, and Japan.

Summary of Background Data. Because clinical decision making becomes more data driven, risk stratification will be crucial to minimize complications. Spine surgeons worldwide face this issue, leading to parallel efforts to address risk stratification. This raises the question as to whether pooled data would be valuable and whether models generated in one country would be applicable to other populations.

Methods. Predictors of SSI and 30-day readmission from three prospective databases (National Neurosurgery Quality and Outcomes Database [N2QOD] N = 2653, DaneSpine N = 1993, Japan Multicenter Spine Database [JAMSD] N = 3798) were determined and compared to identify common or divergent predictive risks.

Results. Predictive variables differed in the three databases, for both readmission and SSI. Factors predictive for hospital readmission were American Society of Anesthesiologists (ASA) grade in N2QOD ($P = 0.013$, odds ratio [OR] 2.08), fusion levels in DaneSpine ($P = 0.005$, OR 1.67), and sex in JAMSD ($P = 0.001$, OR = 2.81). Associated differences in demographics and procedural factors included mean ASA grade (N2QOD = 2.45, JAMSD = 1.72) and fusion levels (N2QOD = 1.39, DaneSpine = 1.52, JAMSD = 1.34). For SSI, sex ($P = 0.000$, OR = 3.30), diabetes ($P = 0.000$, OR = 2.90), and length of stay ($P = 0.000$, OR = 1.02) were predictive in JAMSD. No predictors were identified in N2QOD or DaneSpine.

Conclusion. Predictors of SSI and hospital readmission differ in the United States, Denmark, and Japan, suggesting that risk stratification models may need to be population specific or adjusted. Some differences in measured parameters exist in the three databases analyzed; however, patient and procedure selection also appear to differ and may limit the ability to directly pool data from different regions. Therefore, risk stratification models developed in one country may not be directly applicable to other countries.

Key words: 30-day readmission, risk stratification, surgical site infection.

Level of Evidence: 2

Spine 2017;42:1311–1315

From the [*]Norton Leatherman Spine Center, Louisville, KY; [†]Rygkirurgisk Center Middelfart, Sygehus Lillebælt, Middelfart, Denmark; [‡]Carolina Neurosurgery and Spine, Charlotte, NC; [§]Department of Orthopedic Surgery, Aalborg University Hospital, Aalborg, Denmark; [¶]Spine Unit, Department of Orthopaedic Surgery, University Hospital of Copenhagen, København, Denmark; [||]Department of Orthopaedic Surgery, Nagoya University Graduate School of Medicine, Nagoya, Japan; [**]Department of Orthopaedic Surgery, Keio University School of Medicine, Tokyo, Japan; [††]Department of Orthopedic Surgery, Faculty of Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan; [‡‡]Department of Orthopaedic Surgery, Hamamatsu University School of Medicine, Shizuoka, Japan; [§§]Department of Orthopaedic Surgery, Dokkyo Medical University School of Medicine, Tochigi, Japan; [¶¶]Department of Neurological Surgery, University of California-San Francisco, San Francisco, CA; and [||||]Department of Neurosurgery, University of Virginia, Charlottesville, VA.

Acknowledgment date: July 8, 2016. First revision date: November 14, 2016. Acceptance date: December 14, 2016.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work.

Relevant financial activities outside the submitted work: board membership, consultancy, patents, royalties, grants, employment, payment for lectures.

Address correspondence and reprint requests to Leah Y. Carreon, MD, MSc, Norton Leatherman Spine Center, 210 East Gray St, Suite 900, Louisville, KY 40202; E-mail: leah.carreon@nortonhealthcare.org

DOI: 10.1097/BRS.0000000000002082

The term "risk stratification" has permeated the spine surgery literature over the past several years, but the concept is certainly not new. Identifying those demographic characteristics and comorbidities that are likely to predict complications or clinical outcomes is the standard

evaluation that surgeons perform in the clinic every day. Recently, national registries and larger datasets have facilitated meaningful statistical analysis, and thus a more data-driven approach to this key component of surgical decision making.[1–6]

Beyond the obvious desire to optimize clinical outcome and minimize complications, risk stratification has other applications that are likely to become increasingly relevant to spine surgeons. Particularly in those fields in which more comprehensive datasets are already available, risk stratification can be an important element in cost-effective strategies.[7–9] In addition, because surgeons are more frequently graded by payers, or through public data reporting, risk stratification will be necessary to insure that comparisons accurately reflect case severity and practice profile.

The challenges underlying risk stratification efforts are not unique to any single region or population. Surgeons around the world face similar issues including limited resources and an expectation for reproducible outcomes. Although surgeons have looked to risk stratification as a tool to improve quality, this effort has been limited by the fact that available datasets are often too small or nonspecific for useful risk stratification. One potential solution might be an ability to pool data accrued in different countries, regions, or study groups. This proposition obviously assumes that predisposing risks factors would be similar across the populations being aggregated.

In an effort to test this hypothesis, we compared risk factors for surgical site infection (SSI) and 30-day hospital readmission in three large datasets from the United States (National Neurosurgery Quality and Outcomes Database [N2QOD]),[10,11] Denmark (DaneSpine),[12] and Japan (Japan Multicenter Spine Database [JAMSD]). In particular, we sought to determine whether differences in demographics or in predictor variables might limit the ability to aggregate these datasets. In essence, we sought to determine whether risk stratification might be a universal language.

MATERIALS AND METHODS

Deidentified data were obtained from N2QOD, DaneSpine, and JAMSD databases regarding patients treated with a surgical procedure including lumbar fusion for degenerative disease. These three datasets are national registries from across each of the countries wherein the sites enter data voluntarily. The manner of data collection and cases included have been previously reported for N2QOD[10,11] and DaneSpine.[12] A comprehensive spreadsheet of potential variables was generated. Available data elements, however, varied across the three registries. Requested data points included basic demographics (age at surgery, sex, body mass index [BMI], race), common comorbidities (smoking status, diabetes, depression), and disease characteristics (diagnosis, ambulatory status, symptom duration). Secondary demographic elements were also collected (employment status, insurance status). Perioperative assessment (ASA grade) and surgical data (levels fused, estimated blood loss, operative time, and length of hospital stay [LOS]) were examined (Table 1).

TABLE 1. Datapoints Requested

Diagnosis
Disc herniation
Recurrent disc herniation
Spondylolisthesis (grade 1 or less)
Stenosis
Adjacent segment degeneration
Mechanical collapse disc collapse
Age
Sex
Body mass index
Insurance
Uninsured
Public/government insurance
Private
Smoking
Smoker
Nonsmoker
Diabetic
Diabetic
Nondiabetic
Ambulation
Independent
Needs assistive device
Wheelchair bound
Symptom duration
<3 mo
>3 mo
Employment
Employed currently working, full time
Employed currently working, part time
Employed, not working
Unemployed
On disability due to spine problem
On disability, NOT due to spine problem
On disability, unspecified
Retired
Homemaker
Student
ASA grade
Number of levels fused
None (decompression only)
One
Two
Three
Four
Estimated blood loss
Operative time
Length of hospital stay
Readmission within 30-days postoperative from all causes
No
Yes
If Yes, reason for readmission

Binary logistic regression analysis was performed to identify factors predictive of 30-day readmission and factors predictive of SSI. Primary analysis was performed identifying potential predictor variables within the individual databases. A secondary pooled analysis was performed as well. For the pooled analysis only variables that were available in

TABLE 2. Factors Predictive of 30-Day Readmission From All Causes Across Three Different Databases

Variable	N2QOD		DaneSpine		JMSD	
	P	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio
Diagnosis	0.857	0.97	0.713	0.96	0.442	1.12
Age	0.150	1.03	0.188	0.99	0.245	1.02
Sex	0.666	1.14	0.247	0.74	0.001	2.81
Body mass index	0.507	1.01	0.645	1.01	0.890	0.99
Insurance	0.357	0.73				
Smoking status	0.325	0.58	0.884	1.04	0.330	0.66
Diabetic	0.667	1.16			0.167	1.57
Anxiety	0.912	0.95			0.597	1.36
Depression	0.314	0.64			0.147	2.64
Ambulation	0.072	1.74			0.289	1.29
Symptom duration	0.448	2.20	0.089	0.49	0.836	0.91
Race	0.773	0.93				
Employment status	0.047	0.89	0.041	1.17	0.269	1.06
ASA grade	0.013	2.08			0.622	1.12
Fusion levels	0.110	0.65	0.005	1.67	0.503	1.18
Estimated blood loss	0.083	1.00			0.617	1.00
Operative time	0.271	1.00			0.124	1.00
Length of stay	0.040	1.03	0.919	1.00	0.930	1.00

all the databases were included. In addition, the data source was included in the regression model as the population from which the subjects were sampled may be different across the databases. All analyses were performed using IBM SPSS v21.0 (Somers, NY).

RESULTS

There were 2653 cases from N2QOD, 1993 cases from DaneSpine and 3798 cases from JAMSD. The incidence of 30-day readmission was 1% in JAMSD, 3% in N2QOD, and 7% in DaneSpine ($P = 0.000$), whereas the incidence of SSI was 2% in JAMSD, 1% in N2QOD, and less than 1% in DaneSpine ($P = 0.000$). Predictive variables differed in the three databases, for both readmission and SSI. Factors predictive for hospital readmission (Table 2) were ASA grade in N2QOD ($P = 0.013$, odds ratio [OR] 2.08), fusion levels in DaneSpine ($P = 0.005$, OR 1.67), and sex in JAMSD ($P = 0.001$, OR = 2.81). Associated differences in demographics and procedural factors included mean ASA grade (N2QOD = 2.45, JAMSD = 1.72) and fusion levels (N2QOD = 1.39, DaneSpine = 1.52, JAMSD = 1.34). For SSI, sex ($P = 0.000$, OR = 3.30), diabetes ($P = 0.000$, OR = 2.90), and LOS ($P = 0.000$, OR = 1.02) were predictive in JAMSD. No predictors were identified in N2QOD or DaneSpine (Table 3).

For the pooled analysis, the datasource ($P = 0.000$, OR = 2.72), sex ($P = 0.011$, OR = 1.46), and the employment status ($P = 0.030$, OR = 1.06) were predictive of 30-day readmission (Table 4). LOS ($P = 0.000$, OR = 1.01), sex ($P = 0.001$, OR = 2.13), datasource ($P = 0.000$, OR = 0.35), body mass index ($P = 0.001$, OR = 1.07), and number of fusion levels ($P = 0.003$, $P = 1.48$) were predictive of SSI (Table 5).

DISCUSSION

Risk stratification has become an integral component of efforts to optimize outcomes, minimize complications, and appropriately use available resources in the delivery of spine care.[1,4] Risk stratification modeling has become increasingly common, as more robust datasets facilitate meaningful statistical analysis. Despite substantial progress, the available datasets have often proven insufficient to effectively guide surgical decision making. Administrative databases, although very large, often lack the necessary granularity to provide patient-level guidance.[13] Study group data, although often very detailed, typically includes a smaller population of patients with more complex pathology that may not be generalizable.[14,15] It would therefore be advantageous if datasets from different national registries or study groups could be effectively aggregated.

The findings of the present study suggest that aggregation of study group or registry data across divergent populations may be more complicated than simply pooling the data. We examined two commonly used benchmarks, SSI and 30-day hospital readmission, as observed in three distinct regional datasets. Although the present study focused on variation in data collected in three different countries, a similar comparison among divergent populations within a single geographic region might also be relevant. Importantly, predictors of SSI and 30-day readmission differed in the United States, Denmark, and Japan.

TABLE 3. Factors Predictive of Surgical Site Infections Across Three Different Databases

Variable	N2QOD		DaneSpine		Japan	
	P	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio
Diagnosis	0.726	0.91	0.287	4.08	0.558	0.93
Age	0.336	0.97	0.059	1.26	0.823	1.00
Sex	0.759	1.20	0.991	0.00	0.000	3.30
Body mass index	0.362	1.03	0.078	1.54	0.510	1.02
Insurance	0.622	0.74				
Smoking status	0.362	0.37	0.446	5.66	0.232	0.66
Diabetic	0.670	1.29			0.000	2.90
Anxiety	0.781	1.27			0.554	0.68
Depression	0.408	0.49			0.194	2.35
Ambulation	0.403	1.63			0.288	1.25
Symptom duration	0.997	7.50	0.996	3.36	0.883	1.06
Race	0.127	0.62				
Employment status	0.949	0.99	0.071	3.09	0.528	1.03
ASA grade	0.036	3.49			0.201	0.76
Fusion levels	0.617	1.24	0.352	3.87	0.038	1.45
Estimated blood loss	0.323	1.00			0.570	1.00
Operative time	0.101	1.00			0.121	1.00
Length of stay	0.599	1.02	0.900	1.00	0.000	1.02

Hospital readmission and SSI rates differed in the three study populations. Although differences in SSI rates were statistically significant, the numeric values were very similar (2% *vs* 1% *vs* 1%). Particularly in light of the inability to clearly differentiate between superficial and deep infections, the reported differences in SSI rate may not be clinically relevant. In contrast, the difference in hospital readmission (1% *vs* 3% *vs* 7%) seems real. The unresolved question is whether this disparity represents a difference in complication profile, or a regional bias regarding the role of hospital readmission in the management of equivalent complications.

Predictive variables for hospital readmission were ASA grade in N2QOD, fusion levels in DaneSpine and sex in JAMSD. Predictive variables for SSI were sex, diabetes, and LOS in JAMSD. There were no predictor variables for SSI in N2QOD or DaneSpine. Demographics also differed somewhat across databases with a higher ASA grade in N2QOD *versus* JAMSD (2.45 *vs* 1.72) and progressively more fusion levels reported in JAMSD, N2QOD, and DaneSpine, respectively (1.34 *vs* 1.39 *vs* 1.52).

The demographic differences between these populations may account for some of the variability in predictor variables. Certainly, the Danish and Japanese populations are relatively homogeneous as compared to the heterogeneous US population. There are also significant cultural differences that might affect surgical indications, specific procedure selection, and indications for readmission. These differences raise questions regarding aggregation of the datasets, and also with regard to analysis of the heterogeneous US population as a single entity.

Aggregation of the three datasets generated a large heterogeneous data pool. Analysis was, however, limited by the fact that data elements were not completely consistent across the three registries. Analysis based upon the shared data elements revealed that sex and number of fusion levels were weak predictor variables for SSI and sex was a weak predictor variable for 30-day readmission. The originating database was the strongest predictor variable for both SSI and 30-day readmission, suggesting that differences in the underlying populations outweighed individual patient or surgical variables as a driver of risk.

The primary weakness of the present study is that the three registries did not collect uniform demographic

TABLE 4. Factors Predictive of All-Cause 30-Day Readmission

Rank		*P*	Odds Ratio
1	Database	0.000	2.72
2	Sex	0.011	1.46
3	Employment status	0.030	1.06

TABLE 5. Factors Predictive of Surgical Site Infections

Rank		*P*	Odds Ratio
1	Length of stay	0.000	1.01
2	Sex	0.001	2.13
3	Database	0.000	0.35
4	Body mass index	0.001	1.07
5	Number of fusion levels	0.003	1.48

variables or perioperative data elements. Although there was enough similarity to allow some meaningful comparisons, the inability to conclusively match the groups leaves many questions unanswered. In addition, as registries often fail to clearly define diagnostic indication for surgery, the variability in outcomes or complications related to indication for surgery is difficult to quantify. Also, despite fairly large cohorts in each region, the numbers may still be insufficient to analyze potential predictor variables that occur infrequently.

This comparative analysis of three large databases from the United States, Denmark, and Japan suggests that aggregation of data from different regions will require both more consistent data collection and careful analysis. Despite the challenges, the potential advantages of larger datasets, and thus the ability to effectively compare a broad range of treatment strategies, are very attractive. At a minimum, this will require consensus on diagnostic delineation, key demographics, and a core set of outcome measures.

Ultimately, risk stratification models developed in one country may not be directly applicable to other countries. Although Denmark and Japan have relatively homogenous populations, the United States does not. This raises the additional question as to whether risk stratification efforts within the United States need to be adjusted by geographic region or demographic subgroup.

➢ Key Points

- ❑ Predictors of SSI and 30-day readmission from three prospective databases (N2QOD N = 2653, DaneSpine N = 1993, JAMSD N = 3798) were determined and compared to identify common or divergent predictive risks.
- ❑ Factors predictive for hospital readmission were ASA grade in N2QOD, fusion levels in DaneSpine and sex in JAMSD.
- ❑ Factors predictive of SSI were sex, diabetes, and LOS in JAMSD. No predictors were identified in N2QOD or DaneSpine.
- ❑ Predictors of SSI and hospital readmission differ in the United States, Denmark, and Japan, suggesting that risk stratification models may need to be population specific or adjusted.

References

1. *AHRQ Registries for Evaluating Patient Outcomes: A User's Guide*, 3rd edn. AHRQ Publication No. 13(14)-EHC111 April 2014.
2. American College of Surgeons National Surgical Quality Improvement Program. Available at: http://site.acsnsqip.org/. Accessed February 01, 2016.
3. Khanna R, McDevitt JL, McClendon J, et al. Utility of readmission rates as a quality of care measure and predictors of readmission within 30 days after spinal surgery: a single-center, multivariate analysis. *Spine* 2015;40:1769–74.
4. Kimmell KT, Algattas H, Joynt P, et al. *Risk modeling predicts complication rates for spinal surgery*. 2015;40:1836–41.
5. Martin CT, Pugely AJ, Gao Y, et al. Incidence and risk factors for early wound complications after spinal arthrodesis in children. *Spine* 2014;39:1463–70.
6. Qin C, Kim JY, Hsu WK. Impact of insulin dependence on lumbar surgery outcomes: a NSQIP analysis of 51,277 patients. *Spine* 2016;41:E687–93.
7. Hughes M, Lip GYH. On behalf of the Guideline Development Group for the NICE national clinical guideline for management of atrial fibrillation in primary and secondary care. *Thromb Haemost* 2008;99:295–304.
8. Kanis JA, Borgstrom F, De Laet C, et al. Assessment of fracture risk. *Osteoporos Int* 2005;16:581–9.
9. Sheldon R, O'brien BJ, Glackhouse G, et al., for the Canadian Implantable Defibrillator Study (CIDS) Investigators. Effect of clinical risk stratification on cost-effectiveness of the implantable cardioverter-defibrillator. The Canadian implantable defibrillator study. *Circulation* 2001;104:1622–6.
10. Asher AL, Speroff T, Dittus RS, et al. The national neurosurgery quality and outcomes database (N2QOD): a collaborative North American outcomes registry to advance value-based spine care. *Spine* 2014;39(22 suppl 1):S106–16.
11. McGirt MJ, Speroff T, Dittus RS, et al. The national neurosurgery quality and outcomes database (N2QOD): general overview and pilot-year project description. *Neurosurg Focus* 2013;34:E6.
12. DaneSpine Årsrapport 2015. Available at: http://drksdanespine.dk/wm420129. Accessed December 31, 2015.
13. Abdul-Jabbar A, Takemoto S, Weber MH, et al. Surgical site infection in spinal surgery: description of surgical and patient-based risk factors for postoperative infection using administrative claims data. *Spine (Phila Pa 1976)* 2012;37:1340–5.
14. Lee MJ, Cizik AM, Hamilton D, et al. Predicting surgical site infection after spine surgery: a validated model using a prospective surgical registry. *Spine J* 2014;14:2112–7.
15. Lewkonia P, DiPaola C, Street J. Incidence and risk of delayed surgical site infection following instrumented lumbar spine fusion. *J Clin Neurosci* 2016;23:76–80.