Title:
Practice variation in Japan: A cross-sectional study of patient outcomes and costs in total joint replacement procedures.

Authors:
Denis A. Cortese, MD; Natalie Landman, PhD; Robert K. Smoldt, MBA; Michael J. Joyner, MD; Sachiko Watanabe, RN, MHSA, MAE; Aki Yoshikawa, PhD

Financial support and disclosure:
Virginia C. Piper Charitable Trust (PI: Denis A. Cortese)
Global Health Consulting Japan (Sachiko Watanabe, Aki Yoshikawa)

Reprints and correspondence:
Robert K. Smoldt, MBA
Healthcare Delivery and Policy Program, Arizona State University
13212 East Shea Blvd, Scottsdale, AZ 85259
robert.smoldt@asu.edu
ABSTRACT

Objective: To assess the degree of variability in the quality and cost of care for total joint replacement in Japan and identify potential practice and policy implications for healthcare systems using prospective payment schemes.

Patients and Methods: We examined patient outcomes and expenditures per case for 11,289 total knee replacement and 12,006 total hip replacement cases performed at >500 hospitals in Japan between April 1 and December 31, 2012. To examine the relationship between post-operative complications, average length of stay (LOS), and joint replacement costs, we used a two tailed Pearson correlation where P<.05 was considered significant.

Results: A defining characteristic of our hospital sample was the low volume of procedures performed: ~90% of hospitals did fewer than 70 total joint replacement cases per year. We also found significant variability in length of stay (10x) and per-case expenditures (3x) for total joint replacements. Although there was no correlation between LOS or per-case expenditures and post-operative complication rates, the relationship between LOS and expenditures per case was significant for both procedures (R = 0.89 for TKR cases; R = 0.86 for THR cases; p < 0.01).

Conclusion: There is significant variability in LOS and per case expenditures for total joint replacement in Japan. This is true even for hospitals without post-operative complications. As a result there is a significant correlation between LOS and cost of care, but not complication rates. We hypothesize that variability in outcomes and cost of care is driven in part by the existing healthcare provider reimbursement structure.
INTRODUCTION
It is well known that U.S. spends significantly more per capita and a higher percentage of the Gross Domestic Product (GDP) on healthcare than other developed nations, yet our health outcomes and access to care remain disparate and inconsistent. Moreover, the rapidly rising costs of healthcare delivery are making medical care increasingly unaffordable to the average citizen and threaten the financial viability of the country. Something needs to be done to fix the American healthcare delivery system; it is not sustainable in its current form.

One of the key strategies proposed to help fix our ailing system is the move toward “value- based” healthcare. Despite claims to the contrary, we believe that value can be defined (as patient outcomes/ cost per patient) and it can be measured.1,2,3 And when value is measured, we find significant variability by medical center and by geographic area across the U.S.4,5 In fact, studies that compare U.S. medical centers on both patient outcomes and cost per patient, show two major trends: 1. Wide variability in both patient outcomes and cost per patient over time; 2. Higher per patient spending does not necessarily produce better results.4

In contrast to the U.S., Japan ranks quite positively in many international comparisons of health systems. For example, it ranks highly on mortality outcomes among developed nations,6 and its health cost per capita is low: $3,213 in Japan compared to $4,016 in France and $8,247 in the United States.7 These data suggest that Japan may be a good healthcare system from which U.S. providers and policy makers might be able to learn as we move toward alternative reimbursement schemes to attain higher value. The comparison with Japan is particularly compelling as the majority of healthcare providers in Japan function in a fee-for-serve payment structure – not unlike the experience of the majority of U.S. providers. Moreover, in 2003, Japan introduced the Diagnosis Procedure Combination (DPC), a per-diem graduated payment system that uses a 14-digit classification code to identify major disease categories as well as treatments and procedures, which has some similarities to the Diagnostic Related Group (DRG system) in the U.S.

Given the lower costs and good overall “mortality amenable to healthcare” outcomes reported for Japan as a whole,6 we hypothesized that unlike the U.S., Japanese hospitals will show lower variability in both per patient costs and outcomes. To test our hypothesis we evaluated patient outcomes and per patient expenditures for total joint replacement patients from a representative group of hospitals in Japan that performed total joint procedures. We were
especially interested in the Japanese total joint data because in the United States the total knee replacement aggregate procedure cost is the second highest among procedures paid for by Medicare and third for all payers nationally, thus making total joint replacements an excellent medical category for this study.8

METHODS
All analyses were carried out on data aggregated at the hospital level from acute care facilities in the Global Health Consulting Japan (GHC Japan) hospital discharge database. GHC Japan, a hospital consulting firm, provides a variety of consulting services to over 800 Japanese hospitals paid on the DPC per-diem methodology. Although DPC paid facilities constitute about one fifth of the total number of hospitals in Japan, DPC facilities tend to be the larger hospitals. As a result, about half (51.7%) of all hospital beds are paid under the DPC system.9 In turn, the GHC Japan database spans eight geographical regions and accounts for 54% of all DPC facilities. We identified 549 hospitals in the GHC Japan hospital discharge database that performed at least one total knee replacement (TKR) and 594 hospitals that performed at least one total hip replacement (THR) between April 1 and December 31, 2012. All hospitals where data were collected for less than six months were excluded from our analysis. In addition, all cases of double-joint replacements were excluded, giving us a final sample size of 514 hospitals and 11,289 TKR cases, and 574 hospitals and 12,006 THR cases. For comparisons with prior studies that examined total joint replacement, we annualized the number of total joint cases performed by a given hospital. These “estimated” case numbers for each facility are based on the number of cases observed over the period of data collection and assume a consistent rate of procedures across a 12-month period. TKR and THR cases were analyzed separately.

Quality metrics:
Post-operative complication rate and average length of stay (LOS) were used as primary outcome measures. Post-operative complications included surgical site infections (SSI) [ICD10: T814 or use of post-operative antibiotic]; reoperation; post-op bleeding [ICD10: T810]; post-op bed sores [ICD10: L89 or use of medication for pressure ulcer]; post-op acute respiratory distress syndrome [ICD10: J80]; pulmonary embolism [ICD10: I269]; post-op septicemia [ICD10: A41, A40]; and in-hospital death. Regardless of whether a case had one or multiple instances of a
post-operative complication (such as SSI and a bed sore), it was counted as a single occurrence for our measure, thus giving us the total number of cases with complications. The post-operative complication rate for each facility was then calculated by dividing the total number of cases with complications by the total number of cases performed by the facility during the time period analyzed. Therefore the complication rate for any given cohort of patient in a given hospital was a conservative estimate. LOS for each procedure was calculated across all cases of that procedure performed in a given facility and included hospital days under DPC payment only, i.e., covering only the acute portion of the hospitalization. Although some of the hospitals in our sample have rehabilitation beds, rehabilitation payments are separate from the DPC and thus any days spent in a rehabilitation bed are excluded from our analysis.

Costs:
Finally, we examined both of the above measures in the context of associated expenditures. We calculated the average hospital revenue per case by dividing the total DPC payments (excluding rehabilitation payments) received by a given facility by the total number of cases performed at that facility. Revenue data were converted from ¥ to $USD assuming a ratio of 1 ¥ to $USD0.0099, the exchange rate on January 4, 2014.

Statistics:
Given the very wide range in the number of total joint cases performed by our sample of hospitals and to reduce the impact of outlier hospitals (e.g., a hospital that performed a single TKR or THR procedure in 2012 and the procedure was accompanied by a post-operative complication, resulting in a 100% complication rate), the averages we report here for Japan represent weighted averages across our sample of Japanese hospitals (weighted by the total number of cases performed by each facility during the period of data collection). Descriptive statistics were used to define the central tendency of continuous variables, while ranges describe the variability of the data. To examine the relationship between post-operative complication rate and LOS (Figures 2A and 2B), LOS and per case expenditures (Figures 3A and 3B), as well as between post-operative complication rate and per case expenditures (Figures 4A and 4B), we used a two-tailed Pearson correlation where p<.05 was considered significant. Although we considered the possibility of performing multivariate regression analyses, we felt that the
aggregate nature of the data did not lend itself well to this approach.

**RESULTS**

Tables 1A and 1B provide a summary view of the baselines characteristics of the total joint data from our sample of Japanese hospitals. Between April 1 and December 31, 2012, 11,289 patients underwent single-knee TKR procedures (Table 1A). The weighted average age at treatment for TKR patients was 75 years. There were 12,006 patients who underwent single-hip THR procedures over the same time period. The weighted average age at treatment for THR patients was 66.7 years (Table 1B).

At the facility level, one of the defining characteristics of our DPC hospital sample was the low volume of procedures performed. We estimated that the average annual hospital volume in our sample of Japanese hospitals was 31 cases and 28 cases for TKR and THR, respectively. In fact, ~25% of the hospitals in our sample did fewer than 10 TKR cases per year (or fewer than one case per month) and ~90% of hospitals did fewer than 70 cases per year. We observed a similar pattern with THR, with ~20% of hospitals in our sample performing 5 or fewer THR cases per year and ~90% of hospitals doing 70 or fewer THRs annually. As can be seen in Tables 1A and 1B, we also observed a very wide variability in the volume of total joint procedures across our hospital sample.

**Post-operative complications**

Figure 1A shows the distribution of post-operative complication rates as a function of the number of cases performed at a given facility. It is clear that for this hospital data set, those hospitals that did fewer than 70 TKR cases per year had much higher rates of post-operative complications. This relationship between low hospital procedure volume and higher rate of complications also held true for THR cases (Figure 1B).

**Duration of care**

As shown in Table 1A, we found that the median LOS for TKR in this Japanese data set was 27.9 days and observed wide variability across the hospitals in our sample – with more than a 10-fold difference between the hospital with the lowest LOS (11.5 days) and the hospital with the highest LOS (144.0 days). The median LOS for THR was 27.6 days and we again observed
significant variability, with a nearly 10 fold difference between the hospital with the lowest LOS (12.6 days) and the hospital with the highest LOS (125.0 days). Figures 2A and 2B summarize the relationship between care duration and post-operative complication rates for each procedure. There was limited correlation between the post-operative complication rate and hospital LOS for both TKR and THR procedures. We calculated an $R = 0.23$, $R^2 = 0.05$ ($p < 0.01$) for TKR and an $R=0.12$, $R^2 = 0.01$ ($p < 0.01$) for THR cases. It is most interesting to note that even those hospitals that had no post-operative complication show a range of LOS from a minimum of 11.5 to a maximum of 144 days for TKR; and a range of LOS from a minimum of 12.6 to a maximum of 125 days for THR.

**Healthcare expenditures:**
We also found widespread variability in expenditures per TKR patient in Japan, with average revenue in the most expensive hospital 3.1 times higher than the average revenue in the least expensive hospital. In case of THR, the most expensive hospital was on average 3.5 times more expensive than the least expensive hospital performing THR. We also observed a significant correlation between per case expenditures and length of stay for both total joint procedures (Figures 3A, 3B). For TKR cases, $R = 0.89$, $R^2 = 0.79$ ($p < 0.01$); $R = 0.86$, $R^2 = 0.74$ ($p < 0.01$) for THR cases.

**Quality vs. cost:**
Finally, we examined the relationship between post-operative complications and per-case expenditures. Figures 4A and 4B show that there was no clear relationship between average hospital revenue per case and post-operative complication rate for either total joint procedure. We calculated $R = 0.25$, $R^2 = 0.06$ ($p < 0.01$) for TKR and $R= 0.07$, $R^2 = 0.01$ ($p = 0.08$) for THR cases. Moreover, we found a large variation in revenue per case even for hospitals that had no post-surgical complications –from $14,049 to $43,299. The average revenue per case is designated by the vertical line in Figures 4A and 4B and the horizontal line designates the average post-operative complication rate.

**DISCUSSION**
In this study we found significant variability in the outcomes and expenditures for total joint
replacement procedures in Japan. Although these findings may not be representative of the entire Japanese system (as they are limited to DPC hospitals and total joint procedures), they are nonetheless unexpected given Japan’s international reputation as a “high-value” healthcare system.

We were particularly surprised by the high variability in surgical volumes and the low average number of total joint cases per hospital (31 for TKR and 28 for THR). In comparison, the average TKR volume for a U.S. hospital is 126, while the average THR volume for a U.S. hospital is 60.\(^9\) We also observed the previously established inverse correlation between the volume of procedures performed and patient complications.\(^10\) Although, the correlation between low surgical volumes and higher complications is not unique to Japan, this may be a bigger issue for Japan than for other countries because \(\sim90\%\) of the hospitals in our sample did fewer than 70 total joint cases per year.

Significant variability was also found in the LOS for total joint replacement, with 10+ fold difference between the hospital with the lowest LOS and the highest LOS. Moreover, the median LOS for TKR across all facilities in this data set was 27.9 days, and LOS for THR was 27.6 days, making it a significant outlier internationally. In comparison, the median LOS for both procedures is 3 days in the U.S., 5 days in Canada, and 9 days in the UK.\(^11\)

It is feasible that these differences in LOS are driven at least in part by the demographics of the patients in our hospital sample. Although, this may be true for TKR patients (median age in Japan of 75 years vs. 69 years in the U.S.), the THR patients in our hospital sample were actually slightly younger (median age of 67 years in Japan vs. 69 years in the U.S.).\(^11\) One could also argue that the longer LOS observed in Japanese hospitals is a result of the observed post-operative complications, yet we found very little correlation between the post-operative complication rate of a given hospital and its patient LOS. In fact, the widest range in LOS for both procedures was observed in the subset of hospitals without any post-operative complications.

In contrast, we found a very strong correlation between hospital LOS and the average expenditures per total joint case. This is perhaps not surprising given that Japan’s DPC per-diem payment mechanism creates an incentive to increase patient LOS. Finally, and in a surprising parallel to the U.S., we also found widespread variability in expenditures per total joint patient in Japan, with the average per case revenue in the most expensive hospital more than 3 times higher
than the average per case revenue in the least expensive hospital.

We hypothesize that the variability in value observed both in the U.S. and in Japan is driven at least in part by the payment structure in which healthcare providers operate. Both countries pay providers without any reference to patient outcomes or overall care efficiency. Moreover, the recently introduced DPC graduated per diem payment in Japan incents long lengths of stay,\(^{12}\) while the DRG payment system which includes only hospital (and not physician) services in the U.S. lacks incentives to coordinate care among providers.

We are aware that our research contains several meaningful limitations that hinder the strength of the data analysis and conclusions. Among these limitations, ours was a retrospective, observational study. Further, our research focused on only 2 surgical procedures, and the data for those procedures (when filtered through our study inclusion criteria) represented only 25-35\% of the data for those procedures nationwide in Japan. Finally, it is feasible that in addition to the payment mechanism, other factors (e.g., patient co-morbidities, adequacy of rehabilitation services, ability to transition to home) contributed to the observed variability in LOS and per-case expenditures. Further studies are warranted to address these limitations.

CONCLUSION

We found significant variability in the outcomes (post-operative complication rate and average length of stay) and cost of care for total joint replacement patients in Japan, which may be driven at least in part by the existing healthcare provider reimbursement structure (i.e., a structure that may unintentionally reward excessive practice variability and the associated costs downstream). Given this variability, we believe that, similar to the U.S., it may be possible for Japan to move its health systems to higher value by changing its present payment system. If we want high value healthcare, we are more likely to get it if we actually start paying for it.
REFERENCES


5. Institute of Medicine. New data on geographic variation. 2011, Available at http://www.iom.edu/Activities/HealthServices/GeographicVariation/Data-Resources.aspx


8. AHRQ. Weighted national estimates from HCUP Nationwide Inpatient Sample (NIS), 2011, Agency for Healthcare Research and Quality (AHRQ), based on data collected by individual States and provided to AHRQ by the States. Total number of weighted discharges in the U.S. based on HCUP NIS = 38,590,733.


11. Waddell, J; Johnson, K; Hein, W; Raabe, J; Fitzgerald, G; Turibio, F. Orthopedic Practice

Table 1A: Metrics for Total Knee Replacement

<table>
<thead>
<tr>
<th>Patients (n=11,289)</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>75.0</td>
<td>74.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospitals (n=514)</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection duration(^a) (months)</td>
<td>8.50</td>
<td>9.00</td>
<td>6.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Annual Estimated Hospital Volume(^b)</td>
<td>30.9</td>
<td>20.0</td>
<td>1.33</td>
<td>381</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complication Rate(^c)</td>
<td>4.78%</td>
<td>2.20%</td>
<td>0.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Average Hospital Revenue/TKR case(^d) (dollars)</td>
<td>$19,192</td>
<td>$18,909</td>
<td>$14,049</td>
<td>$43,299</td>
</tr>
<tr>
<td>Length of Stay(^e) (days)</td>
<td>28.9</td>
<td>27.9</td>
<td>11.5</td>
<td>144</td>
</tr>
</tbody>
</table>

\(^a\) Data were collected for hospitals that performed at least one TKR between April and December 2012

\(^b\) Data were collected for a 6-9 month period and annualized in order to give an estimated annual case load.

\(^c\) Post-operative complication rate is calculated for each facility and identified by the presence of one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death.

\(^d\) Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of TKR cases performed at that facility for the duration of observation and weighted by the number of cases each facility performed.

\(^e\) Average length of stay is calculated as a weighted average across all facilities and includes the acute portion of hospitalization only.

All data were aggregated at the level of an individual hospital prior to analysis.
<table>
<thead>
<tr>
<th></th>
<th>Patients (n=12,006)</th>
<th>Hospitals (n=574)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Minimum</td>
</tr>
<tr>
<td><strong>Patients (n=12,006)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>66.7</td>
<td>66.4</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Hospitals (n=574)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collection duration(^a) (months)</td>
<td>8.92</td>
<td>9.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Annual Estimated Hospital Volume(^b)</td>
<td>28.1</td>
<td>14.7</td>
<td>1.33</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication Rate(^c)</td>
<td>5.54%</td>
<td>2.70%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Average Hospital Revenue/THR case(^d) (dollars)</td>
<td>$22,569</td>
<td>$22,204</td>
<td>$16,149</td>
</tr>
<tr>
<td>Length of Stay(^e) (days)</td>
<td>28.9</td>
<td>27.6</td>
<td>12.6</td>
</tr>
</tbody>
</table>

\(^a\) Data were collected for hospitals that performed at least one TKR between April and December 2012.
\(^b\) Data were collected for a 6-9 month period and annualized in order to give an estimated annual case load.
\(^c\) Post-operative complication rate is calculated for each facility and identified by the presence of one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op sepsis, and in-hospital death.
\(^d\) Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of TKR cases performed at that facility for the duration of observation and weighted by the number of cases each facility performed.
\(^e\) Average length of stay is calculated as a weighted average across all facilities and includes the acute portion of hospitalization only.

All data were aggregated at the level of an individual hospital prior to analysis.
Data were collected from hospitals for a 6-9 month period. The number of cases were then annualized in order to give an estimated annual case load. Post-operative complication rate is calculated for each facility and identified by one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death. 514 sample DPC hospitals displayed and each point is representative of one hospital.
Figure 1B: Post-operative complication rates vs. estimated annual hospital THR volume

Data were collected from hospitals for a 6-9 month period. The number of cases were then annualized in order to give an estimated annual case load. Post-operative complication rate is calculated for each facility and identified by one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death. 574 DPC sample hospitals displayed and each point is representative of one hospital.
Figure 2A: TKR average hospital length of stay vs. post-operative complication rates

Average length of stay is calculated across all cases of TKR performed in a given facility and includes the acute portion of hospitalization only. Post-operative complication rate is calculated for each facility and identified by one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death. 514 hospitals displayed and each point is representative of one hospital.

R = 0.23
R² = 0.05
Average length of stay is calculated across all cases of THR performed in a given facility and includes the acute portion of hospitalization only. Post-operative complication rate is calculated for each facility and identified by one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death. 574 DPC sample hospitals displayed and each point is representative of one hospital.
Average length of stay is calculated across all cases of TKR performed in a given facility and includes the acute portion of hospitalization only. Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of cases performed at that facility for the duration of observation. 514 DPC sample hospitals displayed and each point is representative of one hospital.
Average length of stay is calculated across all cases of THR performed in a given facility and includes the acute portion of hospitalization only. Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of cases performed at that facility for the duration of observation. 574 DPC sample hospitals displayed and each point is representative of one hospital.
Post-operative complication rate is calculated for each facility and identified by one or more complications including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op sepsis, and in-hospital death. Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of cases performed at that facility for the duration of observation. Each point represents one hospital.
Post-operative complication rate is calculated for each facility and identified by one or more complication including surgical site infections, reoperation, post-op bleeding, post-op bed sores, post-op acute respiratory distress syndrome, pulmonary embolism, post-op septicemia, and in-hospital death. Average hospital revenue is calculated by dividing the total DPC payments received by a facility by the total number of cases performed at that facility for the duration of observation. Each point represents one hospital.