RESEARCH PAPER

COMPARISON OF BALDRIGE AWARD APPLICANTS AND RECIPIENTS WITH PEER HOSPITALS ON A NATIONAL BALANCED SCORECARD

DAVID A. FOSTER, PHD, MPH
CENTER FOR HEALTHCARE ANALYTICS

JEAN CHENOWETH, SENIOR VICE PRESIDENT
CENTER FOR HEALTHCARE IMPROVEMENT AND 100 TOP HOSPITALS® PROGRAMS

OCTOBER 2011
The Baldrige Board of Governors identified the Thomson Reuters 100 Top Hospitals® national study as a statistical approach for assessing similar aspects of organizational improvement and performance in hospitals and health systems. This research investigates the relationship between healthcare organizations that have achieved recognition through the Malcolm Baldrige National Quality Award process and highest performance (top 3 percent of all hospitals) on the 100 Top Hospitals National Balanced Scorecard.

The results of this study demonstrate a number of strong associations between the performance of Baldrige hospitals (award winners and applicants receiving site visits) and 100 Top Hospitals award winners. The analyses show:

- There is substantial agreement between the results of the Baldrige process and the data-based 100 Top Hospitals award: Hospitals that have been Baldrige award recipients are significantly more likely than their peers to win a 100 Top Hospitals national award.
- Baldrige hospitals were significantly more likely than their peers to display faster five-year performance improvement.
- Baldrige hospitals, as a group, were about 83 percent more likely than non-Baldrige hospitals to be awarded a 100 Top Hospitals national award for excellence in balanced organization-wide performance.
- Baldrige hospitals outperformed non-Baldrige hospitals on nearly all of the individual measures of performance used in the 100 Top Hospitals composite score.

The results demonstrate that hospitals using the Baldrige process exhibit significantly higher rates of improvement in balanced organizational performance than non-Baldrige hospitals. And hospitals using the Baldrige process are significantly more likely than peers to become 100 Top Hospitals award winners, thereby achieving performance equal to or better than the top 3 percent. Although the Baldrige process and the 100 Top Hospitals statistical measurements are quite different, the results of this study suggest that the methods are complementary and identify similarly high-achieving organizations.

INTRODUCTION

The Baldrige Program was established by the United States Congress in 1987 for the purpose of improving the competitiveness and performance of U.S. organizations. In 2001, the Award, originally open only to for-profit businesses, was expanded to include organizations in the education and healthcare sectors.
Assessing the impact the program has on the performance of the not-for-profit and government organizations is challenging because of the absence of normative data and comparisons. The Baldrige program recognizes that many factors contribute to the performance of complex organizations and that demonstrating causal relationships between the use of the Baldrige approach and any particular organization’s success, or lack thereof, would be difficult. More recently, efforts to demonstrate such impact have turned to assessing whether organizations that have demonstrated a certain level of maturity in their use of the Baldrige approach outperform their industry as a whole on key metrics, when such metrics are available.

The healthcare sector is of particular interest to Baldrige program stakeholders. Since 2005, healthcare organizations have represented more than 50 percent of the applicants for the Baldrige Award as healthcare spending approaches 20 percent of the nation’s gross domestic product. In viewing organizational performance, the Baldrige criteria take a balanced approach requiring the evaluation of results across five categories: product and process (healthcare) outcomes, customer-focused outcomes, workforce-focused outcomes, leadership and governance outcomes, and financial and market outcomes. To assess the success of the program in the healthcare sector, the program made an effort to identify a data set that would closely match this view.

The 100 Top Hospitals program was selected because it uses independent public data to measure the overall organizational performance of hospitals and health systems to identify those organizations that set the national benchmarks for delivering balanced excellence (quality, efficiency, financial stability, and patient satisfaction) and high value to the communities served. The program is based on a national balanced scorecard and reflects the impact of leadership on achievement of organizational mission and goals.

To continue the Baldrige commitment to protecting the privacy of individual applications, and to focus on effective users of the Baldrige approach, it was decided to study all hospitals identified by the Baldrige Panel of Judges as potential role model organizations. This included all Baldrige Award winners and all organizations that received site visits in the past five years. The applications of award recipients are in the public domain. Permission to review the applications was received from those reaching the site visit stage, from 2002-2010, without public identification. This group was then compared to the 100 Top Hospitals data sets in an attempt to assess the value created by organizations employing the Baldrige approach.

The 100 Top Hospitals study is annual, quantitative research that identifies the hospitals with the best facility-wide performance. For 18 years, the 100 Top Hospitals research has helped hospital leaders gain an objective comparison of their performance to similar hospitals and develop a balanced plan to reach for excellence. At the heart of this research is the 100 Top Hospitals National Balanced Scorecard, a set of actionable measures that evaluates performance excellence in clinical care, patient perception of care, operational efficiency, and financial stability.

The 100 Top Hospitals study was selected as a basis for comparison because the study’s measurement areas match well with the Baldrige performance categories. This table shows how the two organizations’ performance metrics align:

<table>
<thead>
<tr>
<th>BALDRIGE OUTCOMES</th>
<th>100 TOP HOSPITALS PERFORMANCE CATEGORIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership and governance</td>
<td>Composite score</td>
</tr>
<tr>
<td>Product and process</td>
<td>Patient outcomes: mortality, complications, patient safety, core measures</td>
</tr>
<tr>
<td>Customer focus</td>
<td>HCAHPS score</td>
</tr>
<tr>
<td>Financial and market</td>
<td>Profitability, expenses, length of stay</td>
</tr>
</tbody>
</table>

Both methods show a correlation between strong governance and high performance on a balanced scorecard, a concept that is backed up by research, including that of Dr. Lawrence Prybil.
METHODOLOGY

Data Sources
Baldrige provided a list of 38 hospitals to be included in this study. This list includes nine Baldrige award applicants with site visits since 2007 (site-visit hospitals that gave permission for release of identities) and 29 hospitals that won a Baldrige award between 2002–2010. This group of 38 is referred to collectively as “Baldrige hospitals” throughout this document. Those who won a Baldrige award are called “Baldrige winners.” All other hospitals in the study are assumed to have no Baldrige activity (neither a site visit nor an award) and are referred to as non-Baldrige hospitals, or peers.

Baldrige winners are selected by evaluating results across five categories: product and process (health care) outcomes, customer-focused outcomes, workforce-focused outcomes, leadership and governance outcomes, and financial and market outcomes. They are also evaluated against six process categories.

We used 100 Top Hospitals databases to identify hospitals that have received a 100 Top Hospitals award for current organization-wide performance (100 Top Hospitals national award winners) or the highest five-year rate of performance improvement (100 Top Hospitals: Performance Improvement Leaders5 award).

Thomson Reuters identifies 100 Top Hospitals award winners using public data and a balanced scorecard with a focus on four domains — quality, efficiency, finance, and consumer assessment of care. The study accounts for differences in size and teaching status and uses only publicly available data. We compared Baldrige hospitals to hospitals included in 100 Top Hospitals studies with data from years 2002–2009. The 100 Top Hospitals studies generally include approximately 3,000 hospitals. Among these in-study hospitals, there were a mixture of Baldrige and non-Baldrige hospitals. See the appendix for more information about the 100 Top Hospitals study methodology.

Analysis
To determine associations between the Baldrige and 100 Top Hospitals designations of hospital performance and improvement, we:

• Measured the association between Baldrige hospitals and the overall, organization-level score (composite) on the 100 Top Hospitals performance metric, for data years 2002–2009.
• Measured the association between Baldrige hospitals and individual metrics comprising the 100 Top Hospitals composite performance measures, for data years 2005–2009:
  – Risk-adjusted mortality index (in-hospital)
  – Risk-adjusted complications index
  – Risk-adjusted patient safety index
  – Core measures mean percent
  – Severity-adjusted average length of stay
  – Case mix- and wage-adjusted inpatient expense per discharge
  – Adjusted operating profit margin
• Assessed whether Baldrige award winners are more or less likely to demonstrate faster rates of improvement than peers on the 100 Top Hospitals organization-level composite score and/or the individual performance measures listed above.
• Evaluated whether Baldrige award winners are more or less likely to be 100 Top Hospitals winners than non-Baldrige award winners, for data years 2002–2009.

In the initial analyses, we used descriptive statistics to characterize counts, proportions, and averages. The comparative analyses required adjustment for the 100 Top Hospitals comparison group (small, medium, and large community; teaching, and major teaching hospitals — see appendix for details) because all 100 Top Hospitals rankings and identification of winners are comparison-group specific.

We used regression models in the analytical comparisons to evaluate the likelihood that Baldrige hospitals would also be hospitals that received recognition as a national 100 Top Hospitals award winner or as having significant five-year improvement rates.
Though this study aims to uncover associations between the Baldrige and 100 Top Hospitals designations of hospital performance and improvement, it is not possible to identify causal relationships. In particular, the timing of any associations between Baldrige hospitals and 100 Top Hospitals performance is uncertain. To evaluate this issue, we established lag periods in either direction to characterize the strength of any discovered associations in terms of temporal relationships between Baldrige and 100 Top Hospitals awards.

The unit of analysis for this investigation was the hospital year. These analyses treated each hospital with one year of data as a unit of analysis. When that hospital had a new year of data, it would then represent a new unit of analysis. Note that in any given year of 100 Top Hospitals data, the number of Baldrige hospitals that match with the 100 Top Hospitals in-study hospitals may vary due to standard exclusions that are implemented when producing the 100 Top Hospitals study databases.

We compared the evaluation between Baldrige activity and 100 Top Hospitals performance across 13 time lags, which are defined as the number of years between the Baldrige activity for any given hospital and that hospital’s performance on 100 Top Hospitals. The 13 lag periods range from -6 years, in which the 100 Top Hospitals evaluations occurred six years before the Baldrige activity, to +6 years, in which the 100 Top Hospitals award would or would not have been received six years after the Baldrige activity.

In general, we made the comparisons between Baldrige and 100 Top Hospitals by matching the Baldrige winners and site-visit hospitals, or, for some analyses, just Baldrige winners, with all in-study hospitals included in a given 100 Top Hospitals study. This allowed us to evaluate the proportion of hospitals, either Baldrige or non-Baldrige, that were awarded 100 Top Hospitals status. To adjust for hospital comparison group, these comparisons were implemented through the application of either linear or logistic regression models in which a hospital comparison group was entered as a categorical adjustment variable. The analysis of lag-periods was, as were other analyses, adjusted for a hospital comparison group, but additionally included adjustment for year of data.

RESULTS

Baldrige Award Recipients More Likely to be 100 Top Hospitals Award Winners

When evaluating what proportion of Baldrige winners (excluding site-visit hospitals as defined above) won a 100 Top Hospitals award with data from 2002–2009, we found the following (Figure 1):

• Baldrige hospitals were more than twice as likely as their peers to win a 100 Top Hospitals national award. A full 7 percent of Baldrige winner hospitals also won a 100 Top Hospitals award, while only about 3 percent of their non-Baldrige peers won.
• Baldrige hospitals were more likely than their peers to display top five-year performance improvement. More than 27 percent of Baldrige winner hospitals also won a 100 Top Hospitals: Performance Improvement Leaders (PIL) award, while only 3 percent of their non-Baldrige peers won the award. This association was not seen in previous years of the PIL study. This result was based on the most recent 100 Top Hospitals longitudinal study of performance improvement, which used data years 2005–2009. Earlier 100 Top Hospitals longitudinal, or performance improvement, studies did not show a statistically significant difference between Baldrige winners and non-Baldrige winner hospitals.
• These differences are statistically (and substantively) significant.
• These statistics were developed using the regression model to adjust for hospital size and teaching status (100 Top Hospitals comparison groups — see appendix for details).
Associations Between Baldrige and 100 Top Hospitals Awards: Detailed Results
We performed a lag-time analysis to evaluate whether there was a temporal relationship between achievement in Baldrige and the 100 Top Hospitals study. In other words, we wanted to find whether the data suggested that winning Baldrige most often preceded winning 100 Top Hospitals status, if there was no apparent temporal relationship between Baldrige and 100 Top Hospitals success, or if it most often happened that winning 100 Top Hospitals status preceded winning the Baldrige award. To do this, we compared the proportion of hospitals that did receive a 100 Top Hospitals award among the Baldrige hospitals (winners and site visit hospitals) within given lag periods.

Over the 13 lag periods, Baldrige hospitals showed an overall proportion of winners of 100 Top Hospitals awards of 6.14 percent, whereas the non-Baldrige hospitals were 100 Top Hospitals winners at a rate of 3.35 percent. Therefore, the Baldrige hospitals, as a group, were about 83 percent more likely than non-Baldrige activity hospitals to be awarded the 100 Top Hospitals national (organization-wide performance) award.

In 10 of the 13 lag periods examined, the Baldrige activity hospitals showed a higher proportion of 100 Top Hospitals winners than did the non-Baldrige hospitals. Only the +3 year lag was statistically significant ($p = 0.007$, odds ratio 5.999, 95 percent confidence interval 1.627–22.120), showing that Baldrige activity hospitals were about six times as likely as non-Baldrige activity hospitals to be 100 Top Hospitals winners. The overall pattern is clear — the Baldrige activity hospitals had consistently higher rates of success in the 100 Top Hospitals national study. These findings are shown in Figure 2 and Table 1 (see next page).
TABLE 1: STATISTICS FOR BALDRIGE AND NON-BALDRIGE HOSPITALS IDENTIFIED AS 100 TOP HOSPITALS WINNERS

<table>
<thead>
<tr>
<th>LAG (YEARS BEFORE OR AFTER BALDRIGE ACTIVITY)</th>
<th>BALDRIGE HOSPITALS INCLUDED IN 100 TOP HOSPITALS STUDY DATABASE</th>
<th>NUMBER OF BALDRIGE HOSPITALS THAT WERE 100 TOP HOSPITALS WINNERS</th>
<th>100 TOP HOSPITALS NATIONAL AWARD WINNERS (%)</th>
<th>DIFFERENCE BETWEEN BALDRIGE AND NON-BALDRIGE HOSPITALS (%)</th>
<th>P-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6</td>
<td>10</td>
<td>1</td>
<td>10.00</td>
<td>3.35</td>
<td>2.412</td>
</tr>
<tr>
<td>-5</td>
<td>15</td>
<td>1</td>
<td>6.67</td>
<td>3.35</td>
<td>1.621</td>
</tr>
<tr>
<td>-4</td>
<td>15</td>
<td>1</td>
<td>6.67</td>
<td>3.35</td>
<td>1.493</td>
</tr>
<tr>
<td>-3</td>
<td>18</td>
<td>2</td>
<td>11.11</td>
<td>3.34</td>
<td>2.560</td>
</tr>
<tr>
<td>-2</td>
<td>19</td>
<td>1</td>
<td>5.26</td>
<td>3.35</td>
<td>1.152</td>
</tr>
<tr>
<td>-1</td>
<td>21</td>
<td>2</td>
<td>9.52</td>
<td>3.34</td>
<td>2.241</td>
</tr>
<tr>
<td>0</td>
<td>28</td>
<td>1</td>
<td>3.57</td>
<td>3.35</td>
<td>0.873</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>1</td>
<td>4.76</td>
<td>3.35</td>
<td>1.225</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>0</td>
<td>0.00</td>
<td>3.35</td>
<td>N/A*</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>3</td>
<td>20.00</td>
<td>3.34</td>
<td>5.999</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>1</td>
<td>6.25</td>
<td>3.35</td>
<td>2.677</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>0</td>
<td>0.00</td>
<td>3.35</td>
<td>N/A*</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>0</td>
<td>0.00</td>
<td>3.35</td>
<td>N/A*</td>
</tr>
<tr>
<td>Overall</td>
<td>228</td>
<td>14</td>
<td>6.14</td>
<td>3.35</td>
<td>--</td>
</tr>
</tbody>
</table>

*No Baldrige hospitals won a 100 Top Hospitals national award this year.
Performance Measure-Specific Results With Lag Time
The next phase of the analysis focused on measure-specific results. We made these comparisons at lag year +3 because the analysis of lag times indicated that this time difference showed the largest difference between Baldrige and non-Baldrige hospitals in terms of winning the 100 Top Hospitals award.

To compare the two groups, we placed hospitals in percentiles of performance, then averaged the percentile for each group. With the exception of expense per adjusted discharge, the Baldrige hospitals outperformed non-Baldrige hospitals on all of the 100 Top Hospitals measures used in these study years, although only the Core Measures results were statistically significant. We analyzed both adjusted and non-adjusted data (adjusted for hospital comparison group). Table 2 shows the results.

<table>
<thead>
<tr>
<th>100 Top Hospitals Performance Measure</th>
<th>Percentile Performance (Higher Is Better), Non-Adjusted</th>
<th>P-Value</th>
<th>Balance Versus Non-Baldrige§</th>
<th>Balance Hospitals Performed…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk-Adjusted Mortality Index</td>
<td>56.820</td>
<td>50.010</td>
<td>0.311</td>
<td>7.57% Better</td>
</tr>
<tr>
<td>Risk-Adjusted Complications Index</td>
<td>51.540</td>
<td>50.020</td>
<td>0.861</td>
<td>1.30% Better</td>
</tr>
<tr>
<td>Patient Safety Index</td>
<td>58.230</td>
<td>50.010</td>
<td>0.274</td>
<td>8.17% Better</td>
</tr>
<tr>
<td>Core Measures Score*</td>
<td>88.810</td>
<td>86.320</td>
<td>0.016</td>
<td>4.90 percentage points Better</td>
</tr>
<tr>
<td>Severity-Adjusted Average Length of Stay†</td>
<td>59.280</td>
<td>50.010</td>
<td>0.109</td>
<td>11.69% Better</td>
</tr>
<tr>
<td>Case Mix- and Wage-Adjusted Inpatient Expense*</td>
<td>$6,015</td>
<td>$5,920</td>
<td>0.530</td>
<td>$346 Worse</td>
</tr>
<tr>
<td>Adjusted Operating Profit Margin*</td>
<td>4.787%</td>
<td>3.186%</td>
<td>0.798</td>
<td>2.03 percentage points Better</td>
</tr>
</tbody>
</table>

* Original units reported in CMS data (i.e., not percentile data) because these measures are not normalized by hospital comparison group (bed size and teaching status).
† For this measure, lower is better.
§ Adjusted for hospital comparison group.

LIMITATIONS
Thomson Reuters relies on the accuracy of Medicare administrative claims data and does not attempt to infer or correct values found therein. Hospitals missing data required to calculate 100 Top Hospitals performance measures are excluded from the study, as are:

- Specialty hospitals
- Federally owned hospitals
- Non-U.S. hospitals
- Hospitals with fewer than 25 acute-care beds
- Hospitals with fewer than 100 Medicare patient discharges in the year of study
- Hospitals with Medicare average lengths of stay longer than 25 days in the year of study
- Hospitals with no reported Medicare patient deaths in the year of study
- Hospitals for which a current Medicare Cost Report was not available
- Hospitals with a current Medicare Cost Report not based on a 12-month reporting period

The statistical analyses used here do not take into account the “nesting” or lack of independence within results for a given hospital involving multiple ascertainment of the same hospital over time. In other words, when a hospital contributes multiple hospital years of information based on multiple years of the study — while each of those years is in fact representative of a different sample of patients — technically, the hospital results for multiple years of ascertainment would be more highly correlated across those multiple years than they would be for the results across multiple hospitals. Even so, this limitation would
tend to impact estimates of the variance rather than estimates of the differences in performance between Baldrige and non-Baldrige hospitals.

It should also be noted that in the 100 Top Hospitals database, hospitals that file cost reports jointly with other hospitals under one provider number are analyzed as one organization. Baldrige applicants, on the other hand, can define the number of hospitals included in their application. For this research, we used the application definition.

Another limitation of this investigation is that administrative data are being used. While it is very common practice in the industry to use administrative data, it is acknowledged that such data does not contain the kind of clinically detailed information that would be ideal for risk adjustment. Therefore, any reports of associations here are not meant to imply a causal connection. In other words, we can study correlations or associations here, but cannot make conclusions related to actual cause-and-effect type of relationships.

CONCLUSIONS

Despite difficulties related to sparse data spread over many years, the results of this investigation clearly show that Baldrige hospitals (those that won the Baldrige Award or had site visits related to the award), performed better than similar hospitals (by bed size and teaching status) that were not Baldrige hospitals. Specifically, when we examined the three-year lag between the Baldrige activity and subsequently winning the 100 Top Hospitals national award, we found the Baldrige-activity hospitals to be almost six times as likely to become 100 Top Hospitals winners as non-Baldrige hospitals. The difference was statistically significant.

In examining the measure-specific differences between Baldrige and non-Baldrige hospitals, it is noteworthy that every measure used in the 100 Top Hospitals study except — expense per adjusted discharge — showed better performance by the Baldrige hospitals than for the non-Baldrige hospitals, although only Core Measures showed a statistically significant difference. In other words, while only one measure difference was statistically significant, almost all measures showed the same pattern of better performance by the Baldrige activity hospitals.

Finally, in examining the longitudinal performance using the 100 Top Hospitals Performance Improvement Leadership study for data years 2005–2009, we again see a statistically significantly higher rate of 100 Top Hospitals awards for improvement over time among the Baldrige activity hospitals compared with non-Baldrige hospitals.

Despite severe limitations in sample size, the results of this investigation clearly reveal that hospitals with Baldrige activity outperform similar hospitals without Baldrige activity.
APPENDIX: 100 TOP HOSPITALS, 2011 STUDY METHODOLOGY SUMMARY

The 100 Top Hospitals is a quantitative study that identifies 100 hospitals with the highest achievement on a balanced scorecard based on Norton and Kaplan’s concept with a focus on four domains — quality, efficiency, finance, and consumer assessment of care. The study accounts for differences in size and teaching status and uses only publicly available data. The 100 Top Hospitals studies provide numerous examples of performance excellence, as evidenced in a number of published studies. The four main steps we take in selecting the 100 Top Hospitals are:

1. Building the database of hospitals, including special selection and exclusion criteria. The study focuses on short-term, acute-care, non-federal U.S. hospitals that treat a broad spectrum of patients. The data come from public sources including the Medicare Provider Analysis and Review (MedPAR) dataset, the Centers for Medicare and Medicaid Services (CMS) Hospital Compare dataset, and the Medicare Cost Report. The 2011 study included 2,914 hospitals.

2. Classifying hospitals into comparison groups according to bed size and teaching status. (The number of hospitals included in the 2011 study is in parentheses):
   - Major teaching hospitals (175)
   - Teaching hospitals (435)
   - Large community hospitals (338)
   - Medium community hospitals (1,042)
   - Small community hospitals (924)

3. Scoring hospitals on 10 performance measures:
   - Risk-adjusted mortality index (in-hospital)
   - Risk-adjusted complications index
   - Risk-adjusted patient safety index
   - Core measures mean percent
   - 30-day, risk-adjusted mortality rate for acute myocardial infarction (AMI), heart failure, and pneumonia*
   - 30-day, risk-adjusted readmission rate for AMI, heart failure, and pneumonia*
   - Severity-adjusted average length of stay
   - Case mix- and wage-adjusted inpatient expense per discharge
   - Profitability (adjusted operating profit margin)
   - Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) score (patient rating of overall hospital performance)*

   *Not included in the Baldrige/100 Top Hospitals analysis because results were not available for all data years.

4. Determining the 100 Top Hospitals award winners by ranking hospitals relative to their comparison group. Within the five hospital comparison groups, we ranked hospitals on the basis of their performance on each of the 10 measures relative to other hospitals in their group. We then summed each hospital’s performance-measure rankings and re-ranked, overall, to arrive at a final rank for the hospital. The hospitals with the best final rank in each comparison group were selected as the winners. All measures except the 30-day mortality rate and 30-day readmission rate received a weight of one in the final ranking process. For the 30-day mortality and readmission rate measures, we give the rates for each of the conditions (AMI, heart failure, and pneumonia) a weight of 1/6 in the final 100 Top Hospitals ranking process for winner selection.

For a detailed methodology of the 100 Top Hospitals study, you may download the study abstract at www.100tophospitals.com/top-national-hospitals/.
REFERENCES