Benchmarking in thoracic surgery

Jorge L. Freixinet a,*, Gonzalo Varela b, Laureano Molins c, Juan J. Rivas d, José M. Rodríguez-Paniagua e, Pedro López de Castro f, José M. Izquierdo g, Juan Torres h

a Thoracic Surgery Service, Hospital de Gran Canaria “Dr. Negrín”, C/ Barranco de la Ballena s/n, 35010 Las Palmas de Gran Canaria, Canary Islands, Spain
b Thoracic Surgery Service, Hospital Universitario, Salamanca, Spain
c Thoracic Surgery Service, Hospital Sagrado Corazón, Barcelona, Spain
d Thoracic Surgery Service, Hospital Miguel Servet, Zaragoza, Spain
e Thoracic Surgery Service, Hospital General Universitario, Alicante, Spain
f Thoracic Surgery Service, Hospital Germans Trias i Pujol, Badalona, Barcelona, Spain
g Thoracic Surgery Service, Hospital de Donostia, San Sebastián, Spain
h Thoracic Surgery Service, Hospital Virgen de la Arrixaca, Murcia, Spain

Received 16 May 2010; received in revised form 6 October 2010; accepted 8 October 2010; Available online 27 November 2010

Abstract

Objective: Presentation of an experience in benchmarking in 13 university Spanish thoracic surgery services. Methods: The minimum basic data set (MBDS) for hospitalization, corresponding to 2007, including all registered hospital discharges, was used. The performance of the hospitals was compared using an external reference pattern (SN) and internal average (BMG). Cases were chosen in which a major pulmonary resection (lobectomy or pneumonectomy) was done for bronchogenic carcinoma. Performance indicators were the complexity of the casuistry (average weight and relative weight). Performance results indicators included average length of stay (preoperative, postoperative, and global lengths of stay were analyzed separately for lobectomies as well as pneumonectomies), complications, mortality, and urgent readmissions. Results: A total number of 4778 cases were analyzed, with major thoracic surgeries being prominent with 1779 (37.3%). For average weight, there was a dispersion between 2.5 and 5.68, with an average of 3.45 for the BMG and 3.43 for the SN. There were some very significant differences in morbidity, with groups having a gross rate of few complications (2.6%) up to many (16.1%). The mortality rate ranged between 1.6% and 6.6%. There were considerable differences in urgent readmissions, with gross rates between 2.6% and 7.3%, considering as points of reference 5.4% (BMG) and 4.7% (SN). Concerning the results of pulmonary resections for bronchogenic carcinoma, the index of pneumonectomies was between 8% and 29%. The average length of stay for lobectomy was between 6 and 9.5, with an average of 7 in BMG. In the case of pneumonectomies, it was between 6 and 26 days, with an average of 9 for BMG. Average preoperative stay also varied widely, between 0.2 and 2.4, while postoperative stay was between 7.5 and 12.1. The gross global rate of complications ranged from 2.7% to 36.7%, with points of reference of 15.6% (BMG) and 13.8% (SN). The complication rate ranged from 3% to 33%, with an average of 14.5% for lobectomies, with higher variability for pneumonectomies (0—58%). Conclusions: Benchmarking could be an effective method for improving clinical management. A considerable variability was detected in our study among the participating groups.

© 2010 European Association for Cardio-Thoracic Surgery. Published by Elsevier B.V. All rights reserved.

Keywords: Benchmarking; Thoracic surgery; Bronchogenic carcinoma; Postoperative management; Preoperative management; Morbidity; Mortality; Severity

1. Introduction

Benchmarking (BM) is a process used in business, and deals with comparing efficiency and quality among given products or activities. Its main objective is the attainment of excellence. In the health-care environment, it is being adopted gradually, and BM processes have been described in several medical, health service, and hospital specialties [1,2].

In Spain, BM was carried out for thoracic surgery in 2004, encompassing information from 2002 and 2003 [3]. Nine Spanish thoracic surgery departments participated. In that edition, measures were proposed for the improvement of relevant diagnostic registration in discharge forms, the decrease in inadequate hospital stays, and the standardization of procedures that could establish reliable criteria to improve the quality of pulmonary resection [3]. In 2008, a second edition of BM was done with the participation of 13 Spanish thoracic surgery services. The objective of this study is the presentation of this experience, which could have an important impact on the management of quality parameters in thoracic surgery services.
2. Methods

2.1. Participating centers

Thirteen Spanish thoracic surgery services participated in this study, all located within university hospitals.

2.2. Data source

The information source used was the minimum basic data set (MBDS) for hospitalization, corresponding to 2007. This information was provided directly by the administrative services from discharge reports completed by the physicians who attended each of the patients, without the involvement of any chief of service implicated in the study. Lasist SA, a health information company that specializes in this type of study, did the entire data management process independently and anonymously. As an external reference pattern to compare the performance of the hospitals included in the study, an MBDS hospital database maintained by lasist was used. This database included Thoracic Surgery Department data from 33 Spanish teaching hospitals (the norm). It included more than 16,000 hospitalization episodes from 2006 with the same criteria as those from the BM.

2.3. Case selection

The study cases included all the hospital discharges registered in the participating hospitals during the aforementioned period. For comparative performance assessment, aggregated values of all the participating hospitals were used as an internal reference pattern in this BM study (BMG), excluding three hospitals with incomplete MBDS and the Spanish norm described before as an external reference pattern (SN). Cases were chosen in which a major pulmonary resection (lobectomy or pneumonectomy) was done for bronchogenic carcinoma. The identification of these cases was carried out using codes from the ninth edition of the International Classification of Diseases, Clinical Modification (ICD-9-CM) present in the registered surgical procedures. In cases of pulmonary resection, lobectomy (codes ICE-9-CM: 32.3 and 32.4) and pneumonectomy (codes ICE-9-CM: 32.5 and 32.6).

In order to assess the differences in severity or complexity of cases among included hospitals the average diagnosis-related group (DRG) weight of thoracic surgery cases of each hospital was used (version AP 21). As a comparative measure of complexity among hospitals, the relative weight (ratio of the average weight of study cases of each hospital to the average weight of the norm) was done.

The indicators of performance results, such as average hospital length of stay and readmissions, were adjusted by casemix, using refined DRG, with a subclassification of DRG in categories of severity based on secondary diagnoses registered for each patient.

2.4. Performance indicators

Evaluation of results took into account the following indicators:

1. Average length of stay (LOS). Preoperative, postoperative, and global lengths of stay were analyzed separately for lobectomies as well as pneumonectomies. The average LOS was also evaluated according to severity. The Standard Functional Ratio (SFR) was used as an indicator, a ratio of the observed and expected average LOS (LOS that the hospital would have required if each type of patient had the average LOS for the SN). The indirect standardization method was performed to estimate the expected average LOS, stratifying by refined DRG categories.

2. Mortality. Intra-hospital mortality was evaluated and adjusted for risk. Risk-Adjusted Mortality Index (RAMI) was used as an indicator, resulting from the ratio of the observed and expected mortality in the series of cases included in the study. Expected mortality is calculated through a logistic regression model, which calculates the probability of expected mortality for each patient calculated on a database of more than 3 million Spanish hospital discharges. Independent variables that were included into the model were age, sex, risk of death for first diagnostic code, risk of death for second diagnostic code with maximum risk, risk of death for the procedure with maximum risk, type of admission (urgent/nonurgent), type of DRG (surgical/nonsurgical), level of care provided by the hospital, catchment area (urban/rural), and long-term care transfer policies of the hospital. The performance of this model on the database of pulmonary resection cases is very satisfactory, with an area under receiver operating characteristic (ROC) of 0.93.

3. Complications. A group of 25 general complications and seven sentinel complications were evaluated. The following complications were included: pulmonary complications (atelectasis, pneumonia, and respiratory insufficiency), pleural complications (pneumothorax, empyema, and hemothorax), cardiovascular complications, postoperative hemorrhage, and wound infections. The Risk-Adjusted Complication Index (RACI) was used, resulting from the ratio of the observed number in patients with any of the identified complications in the series of cases included in the study. Expected number of patients with complications was calculated by a logistic regression model similar to that described for mortality. The independent variables included in the logistic model were age, sex, risk of complications of the first diagnostic code, risk of complication of secondary code with maximum risk, risk of procedure with maximum risk, type of admission, type of DRG, level of care, and number of diagnostics for discharge. The performance of this model on the pulmonary resection case database is very satisfactory with an ROC of 0.85.

4. Readmissions. Urgent readmissions before 31 days related to the original admission were analyzed. The Standardized Readmissions Ratio (SRR) was used as an indicator, the ratio of observed and expected readmissions and the norm, adjusted for refined DRG to take the patients’ severity score into account. The indirect standardization method was performed to estimate the expected readmission rate stratifying by refined DRG categories.
2.5. Statistical analysis

Descriptive statistics are presented in this article. No specific 'ad hoc' statistical inferences were expected. As a robust descriptive statistics, the median of hospital values and the standard deviation were used to describe the main variables, as there were only 13 hospitals participating in this experience.

3. Results

A total number of 4778 cases were analyzed, with major thoracic surgeries being prominent with 1779 (37.3%). Pneumothoracies were 10.9% (n = 520) and other thoracic interventions did not exceed 11.5% (n = 548). Between 70% and 90% of the hospital discharges were done by the Thoracic Surgery Service, although in a few cases the discharges were directed to other services (oncology, internal medicine, pneumonology, and others).

3.1. All thoracic surgery cases

A description of the main variables analyzed is shown in Table 1. The median number of cases by hospital was 388 (range: 133–662). The median of the age mean of each hospital case was 56 (range: 56–61 years). The median of the percentage of males in each hospital set of cases was 70%.

The median of the percentage of hospital admissions through the Emergency Department was 38%, ranging from 27% to 46% (see Table 1). The median of mean DRG weights of each hospital case group (AP 21) as shown was 3.34, ranging from 2.57 to 5.68 (see Table 1). Fig. 1 shows the differences in the relative DRG weight of each hospital related to the mean DRG weight of the study group as a measure of complexity. The hospital with the highest complexity was 65% more complex than the mean BMG; the hospital with the lowest complexity was 25% less complex than the mean BMG.

The median of the mean LOS of the BMG was 9.3 days, ranging from 6.9 to 14.6 days. Of these 9.3 days as a median figure, 2.2 were preoperative LOS and 7.5 postoperative (see Table 1). The mean LOS of the SN was 10.7 days. Fig. 2 shows the differences in the standardized LOS by refined DRG (AP 21) of each hospital related to the mean DRG weight of the study group.

The median mortality during the surgery admission episode of all thoracic surgery cases in each hospital of the BMG was 3.4%, ranging from 1.6% to 6.6% (see Table 1). The SN showed a mortality rate of 4.3%. Fig. 3 shows the differences in the RAMI for each hospital. The hospital with the highest value showed 48% more deaths than expected; the hospital with the lowest value showed 55% fewer deaths than expected.

### Table 1. Description of the main variables analyzed.

<table>
<thead>
<tr>
<th></th>
<th>All thoracic surgery cases</th>
<th>Lobectomies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Min</td>
</tr>
<tr>
<td>No. of cases/hospital</td>
<td>388</td>
<td>133</td>
</tr>
<tr>
<td>Age (years)</td>
<td>56</td>
<td>52</td>
</tr>
<tr>
<td>Male gender (%)</td>
<td>70</td>
<td>62</td>
</tr>
<tr>
<td>Admission through Emergency Department (%)</td>
<td>38</td>
<td>27</td>
</tr>
<tr>
<td>DRG weight (AP 21.0)</td>
<td>3.34</td>
<td>2.57</td>
</tr>
<tr>
<td>Length of stay (days)</td>
<td>9.3</td>
<td>6.9</td>
</tr>
<tr>
<td>Preoperative LOS (%)</td>
<td>2.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Postoperative LOS (%)</td>
<td>7.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Mortality during surgery admission (%)</td>
<td>3.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Complications (%)</td>
<td>8.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Postoperative pulmonary complications (%)</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>Postoperative septicemia, abscess or wound infection (%)</td>
<td>1.1</td>
<td>0.2</td>
</tr>
<tr>
<td>1.5</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Complications after procedures related to other organic systems (%)</td>
<td>1.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Urgent readmissions at 30 days (%)</td>
<td>5.5</td>
<td>2.6</td>
</tr>
</tbody>
</table>

* Reported number of cases with preoperative data: 72% in all thoracic surgery cases, 96% in lobectomies.
The median percentage of complications analyzed in this study registered during the surgery admission episode of all thoracic surgery cases in each of the 13 hospitals included into the BMG was 8.9%, ranging from 2.6% to 16.1% (see Table 1). The SN showed a 7.5% rate of patients with complications. Fig. 4 shows the differences in the RACI for each hospital. The hospital with the highest value showed 104% more patients with complications than expected; the hospital with the lowest value showed 36% fewer patients with complications than expected.

The percentage of urgent readmissions at 30 days after all thoracic surgery cases in each of the 13 hospitals included into the study was 5.5%, ranging from 2.6% to 7.4% (see Table 1). The SN showed a percentage of 4.7% urgent readmissions at 30 days. Fig. 5 shows the differences in SRR of each hospital. The hospital with the highest value showed 57% more urgent readmissions at 30 days than expected; the hospital with the lowest value showed 52% fewer urgent readmissions at 30 days than expected.

3.2. Results of pulmonary resections for bronchogenic carcinoma

Table 1 shows the results of the main variables analyzed in this study for the set of lobectomy and pneumonectomy cases. The median number of lobectomies by hospital was 73, ranging from 30 to 142. The median number of pneumonectomies by hospital was 13, ranging from 7 to 30. Only a limited analysis can be made by hospital, given the low numbers for most of the participants, especially in pneumonectomies.

The main performance indicators are described. The median value for the average LOS in each hospital for lobectomy was 9.9 days, ranging between 7.0 and 12.8 days. The median value for the average LOS in each hospital for pneumonectomy was 11.3 days, ranging between 7.1 and 26.5 days (see Table 1).

The median mortality during the surgery admission episode of lobectomy in each of the 13 hospitals included into the BMG was 2.7%, ranging from 0.0% to 7.1% (see Table 1). The median mortality during the surgery admission episode of pneumonectomy in each of the 13 hospitals included in the study was 7.7%, ranging from 0.0% to 17.6% (see Table 1).

The median percentage of complications analyzed in this study registered during the surgery admission episode of lobectomy in each of the 13 hospitals included in the BMG was 12.3%, ranging from 3.3% to 32.9% (see Table 1). The median percentage of complications analyzed in this study that were registered during the surgery admission episode of pneumonectomy in each of the 13 hospitals included in the BMG was 15.4%, ranging from 0.0% to 58.3% (see Table 1).

The urgent readmissions at 30 days were not analyzed due to the low number of events registered.

4. Discussion

The possibility of improving the results of clinical management involves gathering information from obtained data, which has spurred initiatives for massive use of databases, which has shown improvement in clinical results [4]. Initiatives for the comparison of results are recent. Spain had its first experience with BM in thoracic surgery in 2004 [3]. This article is the second experience with these initiatives.

The experience recounted in this article is based on the analysis of a database maintained basically for administrative purposes, which gathers clinical data in release forms filled out by personnel responsible for each patient. The use of large administrative databases for the evaluation of clinical results and for the improvement of the quality of medical assistance has been criticized [5], as these databases do not contain enough clinical information to arrive at conclusions that take into account the quality of health care. Other authors, on the other hand, have defended their use [6,7], and have emphasized that administrative databases are valid for the evaluation of results of health-care attention as long as they meet certain conditions, which include that the measured results (mortality, complications, etc.) are frequent [8]. Despite existing limitations, and given the presence of this database in every hospital and the consolidation extant in many countries, it is efficacious to use it to identify areas for improvement in aspects of clinical governance that do not require great precision in diagnostic selection and registration of clinical information.
From a scientific point of view, for the evaluation of the quality of care in thoracic surgery, it would be better to use specifically designed databases. Future exploitation of current large databases [9,10] will be very advantageous for the improvement of quality in thoracic surgery practice.

BM clinical unit studies should not be considered as a ‘better or worse’ classification of participating units. Its philosophy is based on the comparison of clinical management data without economic consideration. A primary objective is a better comprehension of the internal organization itself with respect to the practices and processes analyzed, given the difficulty in determining differences among proposed systems for the different participating centers without an adequate knowledge of them. Another aim of this study is to obtain an improvement of the systems and results based on learning from units with the best results. However, there are no data from well-designed studies showing beneficial clinical effects from this type of practice [3].

Among the limitations of this study, what stands out is that an evaluation of results is done without analyzing the related processes. It therefore fails to fulfill one of the conditions cited in the literature for the validity of using administrative databases in the evaluation of assistance quality [8]. Although some published evaluations of initial experiences in thoracic surgery processes can be found [11], evaluation of assistance processes in our specialty needs to be further developed and is beyond the objectives of our publication. Another important limitation is the irregular geographic distribution of participating units, the majority in the north central area of Spain and few in the south (only three). In addition, there is the continuing problem that the origin of data is from MBDS codification, with defects highlighted in previous publications [3]. The creation of prospective databases, with internal quality controls, an initiative now underway by the European Society of Thoracic Surgery [11], is of great importance. Another issue for improvement is the questionable quality of data provided by some institutions, which has caused the exclusion of their results in some of the parameters and in the average of the participating groups in the BM.

There are also some notable advantages that encourage the pursuit of BM. The guarantee that the data has not been manipulated makes the study more reliable, and so is the creation of a culture of improvement and analysis of a unit’s own results and then comparing them to the results of other units.

With respect to the data studied in this edition of BM, the measurement of complexity and risk adjustment stands out, which makes the obtained data comparable, avoiding the disparity of the populations that attend the participating centers. The comparison established with the averages of all the participants of the BMG and the SN is also very important. These data have allowed each unit to carry out an internal analysis of their results and plan a more rigorous improvement. The measurement of severity is conditioned upon the quality of discharge reports, which should also include patient comorbidity and all complications produced; these should be taken as a reflection of the type of services included in the discharge forms. The recommendation for improvement and systemization was one of the most important matters proposed by the study.

The figures presented show a notable disparity among the participating units. The volume of discharges varied between 133 and 672, which indicates that many units significantly below and above the norm (n = 477) took part in the study. This could explain the differences in other data, such as the profile of the casuistic and the distribution of hospital discharges by service. The discharges were very irregularly distributed, between those of the specialization and those attributed to others, mainly the pneumonology and internal medicine services. As to the distribution of hospital discharges by DRG, there were units with 100% surgery discharges, and, at the other extreme, units with 60% — majority were between 70% and 75% surgical DRG, the most homogeneous data.

One of the most important matters that influence clinical management of the units is emergency admissions. Among participating units, there were a significant number of admissions of this type, ranging between 28% and 46%. The matter of nonprogrammed admissions has been widely debated, as it is considered that a significant number of emergency admissions imply greater difficulty in good clinical management. This section is one of the others that are considered a candidate for improvement in many units.

An average diagnosis by episode is a very important measurement, which seriously influences the codification of processes and the measurement of severity of the same. The majority of participating services was above the norm [4,6] or very close. It is a good-quality indicator of the participating units and could probably have been influenced by the first edition of BM.

Analysis of mortality has provided data by means of BMG that are within average figures that mark international standards. Its variability among distinct groups has, however, been very significant, especially in pulmonary resection. This is objective data because it does not depend on discharge reports from the services, but rather on data from MBDS provided by the hospitals. Some of the data have provoked measures of revision and control of the procedures for some of the participating groups.

Average hospital stay is a heavily studied parameter in clinical management. It was very variable in our study, coinciding with previous studies. There were stays for pulmonary resections that reached between 7.6 and 13.4, with a BMG average of 10.6. Independent of the introduction of minimally invasive and fast-track measures [12], the standardization and production of clinical routes for pulmonary resection could be ways to improve this parameter, which is tied to better cost effectiveness. BM is therefore a measure of results that are joined to other more generic results of analysis of assistance indicators [11] and patients’ perceived quality [13], and, will therefore, allow for a better evaluation of the activity.

5. Conclusions

BM is an effective method to improve clinical management based on results. A considerable variability was detected in our study among the participating groups. Some of the parameters can be improved by actions such as reducing preoperative LOS, making use of high-resolution outpatient visits to avoid admissions for diagnostic procedures, setting programmed
admissions protocols through admission services, and improving discharge reports and coding to include all existing comorbidity and complications that arise to better describe the patient-care process. It is vital to establish clinical pathways and register the key data in databases to be able to analyze and evaluate performance results.

References


Appendix A. Participant hospitals

- Hospital Clínico y Provincial, Barcelona, Spain: Dr Marco A. Callejas and Dr José M. Gimferrer.
- Hospital Universitario, Salamanca, Spain: Dr Gonzalo Varela.
- Hospital Germans Trias i Pujol, Badalona, Barcelona, Spain: Dr Pedro López de Castro and Dr Emiliano Astudillo.
- Hospital Sagrado Corazón, Barcelona, Spain: Dr Laureano Molins and Dr Juan J. Fibla.
- Hospital Clínico San Carlos, Madrid, Spain: Dr Florentino Hernandez-Tranco.
- Hospital de Gran Canaria Dr Negrín, Las Palmas de Gran Canaria, Spain: Dr Jorge L. Freixinet and Dr Pedro M. Rodríguez.
- Hospital de Cruces, Baracaldo, Spain: Dr Joaquín Pac.
- Hospital Miguel Servet, Zaragoza, Spain: Dr Juan J. Rivas and Dr Primitivo Martinez.
- Hospital Virgen de la Arrixaca, Murcia, Spain: Dr Juan Torres and Dra María J. Roca.
- Hospital Son Dureta, Palma de Mallorca, Spain: Dr Carlos Montero.
- Hospital de Donostia, San Sebastián, Spain: Dr José M. Izuquierdo and Dr Carlos Hernández.
- Hospital General Universitario, Alicante, Spain: Dr José M. Rodríguez-Paniagua and Dr Sergio Bolufer.
- Hospital Juan XXIII, Tarragona, Spain: Dr Emilio Canalis and Dr José Cerón.