Using length of stay and inactive days in the hospital to assess appropriateness of utilisation in Barcelona, Spain

Jordi Alonso, Alvaro Muñoz, Josep M. Antó and the Appropriateness Evaluation Protocol Group of the Hospital Universitari del Mar

Abstract

Study objective – To compare the level of inappropriate utilisation of a teaching hospital in two different calendar years and to analyse the relationship between changes in appropriateness of utilisation and changes in average length of stay.


Setting – Hospital Universitari del Mar, a teaching hospital with 440 beds located in Barcelona, Catalonia, Spain.

Participants – Medical records were randomly selected from records of adults discharged from hospital with a principal diagnosis other than normal delivery or any psychiatric condition. Altogether 750 records with complete data were reviewed for 1988 and 633 for 1990.

Measurements and main results – The appropriateness evaluation protocol (AEP) was used to assess whether or not the admission and each day of the hospital stay were appropriate. For data analysis we used the extension of logistic regression that allows quantification of with-inpatient clustering of inactive days. The average length of stay (LOS) was 11.7 days in 1988 and 9.5 in 1990 (p<0.001). In contrast, the proportion of inappropriate admissions was 12% in 1988 and 19% in 1990 (p<0.001). Among those patients who were appropriately admitted, the average odds of a given day being inactive in 1990 was no lower (OR = 1.09) than the odds in 1988, but the clustering of inactive days was significantly (p<0.001) lower in 1990.

Conclusions – In the hospital studied, a lower average length of stay was not associated with an improved level of appropriate utilisation. Policies exclusively focused on lowering LOS may not directly lead to a reduction in inappropriate hospital utilisation.

(J Epidemiol Community Health 1996;50:196–201)

In an environment of health cost containment, policies targeted at reducing the length of stay in acute hospitals represent a straightforward strategy for lowering expenditure. A number of studies have shown that some hospital care is inappropriate or not urgent, either because it could have been provided at a lower cost (eg, as an outpatient) or because no relevant clinical action was taken during the admission.

In order to increase hospital efficiency, managers in the United States and in Europe are devoting considerable effort to reducing the average length of stay (LOS) in hospital. In the United States, these reductions have been clearly stimulated by new reimbursement systems. In general, prospective payment systems have been associated with shorter hospital stays and better quality of inpatient care. Nevertheless, data indicating a worse outcome at discharge associated with a reduction in LOS after the implementation of prospective payment schemes have also been reported. There is still no definite evidence that a reduction in LOS is associated with a parallel decrease in the number of days spent inappropriately in hospital.

In Catalonia, Spain, where there is a national health system, hospital care is provided by public or privately owned hospitals under agreement with the public administration. Since the mid 1980s, reimbursemnt for some hospitals has been negotiated with the administration based on the previous year’s overall average LOS. Because these hospitals are prepaid for a prespecified length of stay, the system stimulates generic or diagnostic-related group (DRG) reductions in the LOS, since a saving or profit can be made when LOS is less than that for which they are paid.

The specific objectives of the present study were to compare the level of inappropriate bed utilisation in a Spanish teaching hospital in the two years, 1988 and 1990, and to analyse the relationship between changes in the appropriateness of bed utilisation and changes in the average LOS.

Methods

The Hospital Universitari del Mar is a teaching hospital with 440 medical and surgical beds. It is the public referral hospital for two districts (districts 1 and 10) of the city of Barcelona and covers a population of approximately 350,000 people. The occupancy rate ranged from 81% to 88% between 1988 to 1991. In the year 1988, a research project was designed to assess periodically hospital utilisation using the appropriateness evaluation protocol (AEP). Because the research project was descriptive in nature no feedback intervention to hospital physicians or administrators was attempted.
ADAPTATION OF THE AEP
The AEP was developed by Gertman and Restuccia in 1981.\(^2\) It includes a set of explicit criteria for evaluating the appropriateness of the admission based on the clinical status of the patient and the level of services provided. In addition, it provides a set of criteria for evaluating the appropriateness of a given day of hospitalisation based on medical services and nursing care provided and on the clinical status of the patient. It has been used in a number of studies in the USA\(^{1,10}\) and other countries\(^{11,12}\) and has been applied to different patient populations (paediatrics,\(^{13}\) and medical/surgical\(^{14}\) among others).

For the present study, the 1986 revised version of the protocol was used.\(^15\) The adaptation of the AEP and evidence of its validity in Spain have been presented elsewhere.\(^16\) Minimal modifications were introduced by a clinical-epidemiological panel to adapt the instrument to some protocolised interventions in the hospital (see Appendix 1). Results of that study\(^16\) also showed that information in medical records was insufficient to assess appropriateness, and that performance indicators of reviewers (eg, reliability and proportion of overrides used) were satisfactory.

DATA COLLECTION
All discharge records in the hospital’s computerised file corresponding to adults (aged 14 years or more) for whom the principal diagnosis was other than normal delivery or any psychiatric condition were considered eligible for the study. From this sampling frame, a random sample of 835 individuals discharged between January 1 and December 31, 1988 (out of a total of 9047 eligible, sampling fraction = 9.2%) and a random sample of 699 individuals discharged during 1990 (out of a total of 12 899 eligible, sampling fraction = 5.4%) were selected. Data were collected in two different time periods. The change in the sampling fraction between 1988 and 1990 was caused by a reduction in the study resources.

Two physicians with experience in medical abstracting procedures reviewed the 1988 discharges, and two years later a second review was carried out by two other physicians who had formal training in medical abstraction. All four reviewers were trained by one of the co-authors (J Bañeres) and evaluated by reviewing a number of medical records (not included in the final samples). In both periods, actual data collection and abstraction did not start until the performance of the reviewers had been judged satisfactory (including high inter-rater reliability).

For every admission, the following information was systematically reviewed in the medical records: emergency room report, main medical history, clinical course, and nursing record. Of the 835 randomly selected admissions in 1988, 29 medical records (3.5%) could not be located and 28 (3.4%) contained incomplete information. Respective figures for 1990 (n=699) were: 35 (7%) and 12 (1.7%). In all, 778 admissions were assessed for 1988 and 652 were assessed for 1990.

APPRIORIENESS ASSESSMENT
Inappropriate admissions were those in which none of the 16 criteria related to the patient’s condition and to the intensity of care included in the AEP list were met. Similarly, inappropriate days of care were those for which none of the 27 criteria were met; and are denominated herein, inactive days. Every day of care for the admissions included in the study was evaluated (see below; statistical analysis). Overrides (extraordinary criteria) were used when necessary by the reviewers either to classify an admission or a day of care as appropriate or inappropriate, overriding the AEP criteria. Overrides were used in 2-3% of the admissions and in 7-7% of the days of care in the 1988 sample, and in 2-8% of the admissions and 6-4% of the days in 1990 data.

An analysis of the agreement between reviewers was carried out by having medical records reviewed by the two reviewers in each year. For reviewers of 1988 admissions, overall agreement was 88% for admissions and 93% for days of stay; specific agreement values (defined as the proportion of inappropriate admissions or days - among those admissions or days - judged to be inappropriate by at least one of the two reviewers)\(^17\) were 60% and 84%, respectively. For reviewers of 1990 admissions, overall agreement was 96% for admission and 93% for days of stay; specific agreement values were 80% and 70% respectively.

STATISTICAL ANALYSIS
Comparison of the length of stays (LOS) in 1988 and 1990 was carried out by standard non-parametric tests.

Inappropriateness of utilisation was assessed using three different outcomes. Firstly, we considered whether or not the admission was appropriate. In this case, we used the exact Mantel–Haenszel procedure to compare the rates of inappropriate admissions in the two periods after stratifying by gender, age, and hospital ward. The second outcome considered was whether the stay contained at least one inactive day; and we examined the correlation between the first two outcomes. The third outcome was the number of inactive days during the LOS of each patient. The analysis of this outcome required the extension of ordinary logistic regression to incorporate the putative within-patient clustering of inactive days.\(^18\)

The key component of logistic regression methods for clustered data is to allow for the variance of the ordinary binomial distribution to be higher than nominal so that the probabilities of fully appropriate stays (number of inactive days = 0) and of fully inappropriate stays (number of inactive days = LOS) are higher than those of the ordinary binomial distribution. The antilog of regression coefficients can be interpreted as the odds ratios of average probabilities of a given day being inactive during a hospital stay. This method is able to measure clustering of inactivity in stays containing at least one inactive day. Specifically, a scale parameter quantifies the within-patient clustering of inactive days. Software to fit these models includes...
Table 1 Characteristics of the two samples (1988 and 1990) of admissions in the Hospital Universitari del Mar, Barcelona, Spain

<table>
<thead>
<tr>
<th>Variables</th>
<th>1988</th>
<th>1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>750</td>
<td>633</td>
</tr>
<tr>
<td>Gender:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>46%</td>
<td>50%</td>
</tr>
<tr>
<td>Female</td>
<td>53%</td>
<td>49%</td>
</tr>
<tr>
<td>Mean (SD) age (y):</td>
<td>54-4 (19-7)</td>
<td>55-3 (19-9)</td>
</tr>
<tr>
<td>14-28</td>
<td>28%</td>
<td>25%</td>
</tr>
<tr>
<td>40-49</td>
<td>11%</td>
<td>11%</td>
</tr>
<tr>
<td>50-59</td>
<td>15%</td>
<td>15%</td>
</tr>
<tr>
<td>60-69</td>
<td>20%</td>
<td>21%</td>
</tr>
<tr>
<td>70+</td>
<td>24%</td>
<td>26%</td>
</tr>
<tr>
<td>Hospital ward:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>45%</td>
<td>40%</td>
</tr>
<tr>
<td>Internal Medicine</td>
<td>11%</td>
<td>9%</td>
</tr>
<tr>
<td>Cardiology</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Pneumology</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>Others, medicine</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OthGyn</td>
<td>15%</td>
<td>9%</td>
</tr>
<tr>
<td>Orthopedics</td>
<td>12%</td>
<td>13%</td>
</tr>
<tr>
<td>General surgery</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>Urology</td>
<td>2.5%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Others, surgery</td>
<td>10%</td>
<td>11%</td>
</tr>
<tr>
<td>Others</td>
<td>21%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Mean (SD) length of stay (d):</td>
<td>11.7 (12.1)</td>
<td>9.5 (10.5)</td>
</tr>
<tr>
<td>1</td>
<td>4.6%</td>
<td>12.0%</td>
</tr>
<tr>
<td>2</td>
<td>20.0%</td>
<td>18.0%</td>
</tr>
<tr>
<td>4-7</td>
<td>22%</td>
<td>26.2%</td>
</tr>
<tr>
<td>8-14</td>
<td>26%</td>
<td>25.0%</td>
</tr>
<tr>
<td>15+</td>
<td>27%</td>
<td>18.0%</td>
</tr>
</tbody>
</table>

The E G R E T package, the manual for which contains a clear heuristic presentation of logistic regression for clustered data. To facilitate the interpretation and assessment of the clustering of inactive days in the two years studied, we depicted the expected probabilities of the number of inactive days for admissions with three different LOS: 3, 7, and 14 days of stay.

Results
The numbers of patients with complete data in the three outcomes of interest were 750 (96%) and 633 (97%) in the years 1988 and 1990, respectively.

Table 1 shows the characteristics of the 1988 and 1990 samples of patients admitted to the hospital in these two periods. Patients were similar with regard to gender and age at admission. There were differences in the distribution of admissions in different hospital wards: 45.1% were admitted to the medical wards in 1988 compared with 40.6% in 1990 (p = 0.03). There was a very (p<0.001) significant difference in the LOS values. The mean LOS was approximately two days shorter in 1990 (mean = 9.5 days) than in 1988 (mean = 11.7). The percentages of the LOS presented in table 1 indicate that the full distribution of LOS shifted towards shorter stays in 1990.

The first outcome analysed was whether or not the hospital admission was appropriate. The characterisation of the admissions does not take into account future events during the stay, including the length and the inactivity. The proportions of patients whose admission was considered inappropriate were 12% in 1988 and 19% in 1990 (p=0.001) (table 2). The proportion of inappropriate admissions was higher for the 1990 sample in all strata according to gender, age, and hospital ward.

The second outcome analysed was whether or not the stay contained at least one inactive day. This approach reduces the information in the number of inactive days to a zero/one variable according to whether the number of inactive days is zero or greater than zero. Table 3 shows the proportion of admissions with at least one inactive day, according to the LOS and the appropriateness of the admission. In both years, almost all (>95%) the individuals who had been inappropriately admitted to the hospital had at least one inactive day during their stay. In contrast, among those patients appropriately admitted, the higher the LOS, the greater the proportion with at least one inactive day. For every LOS considered, this proportion was higher in 1990, reaching a borderline significance value for stays between 8 and 14 days and being very significant (p = 0.01) for LOS greater than 14 days.

In the third analysis, the dependent variable was the number of inactive days and the independent variables were gender, age, ward of admission, and calendar year as explanatory variables. Because of the collinearity between inappropriate admission and the presence of at least one inactive day during the stay, we restricted this analysis to those patients who had been appropriately admitted (n = 1164). Gender, hospital ward, and year were categorical variables; age was used as a continuous variable and odds were presented corresponding to an increment of 10 years of age (eg, 60 vs 50 years of age). Because LOS was very skewed to the right, logarithmic transformation was used. The results are presented in table 4.

The odds of a given day being inactive were significantly lower in surgical (OR=0.71) and other wards (OR=0.47) compared with medicine. A longer LOS was significantly associated with the number of inactive days (OR=1.53 for twice the length). In contrast with the anticipated hypothesis that a lower average LOS would be associated with a lower number of inactive days, a higher but non-significant probability (OR=1.09, p=0.0462) of inactive days was found for admissions in year 1990 compared with those in year 1988. Females and younger patients had lower probabilities of inactive days but neither of the associations was statistically significant.

There was a significant clustering of inactive days, indicating that when patients have an inactive day the chances are that they present
with more than one. This trend of accumulating inactive days in individuals was significant for both years (p<0.001), but the negative sign of the scale coefficient of year indicates that in 1990 the magnitude was significantly lower than in 1988.

The figure illustrates the model given in table 4. The figure shows the predicted probability of having a different number of inactive days (from 0 to all the stay) for a reference individual (ie, male, age = 54, admitted to medicine) whose length of stay was 3, 7, or 14 days in 1988 and in 1990. For each of the stays, both the probability of having a fully appropriate stay (0 inactive days) or a fully inappropriate stay was higher in 1988. The probability of having a number of inactive days between the above extremes was higher in 1990.

In addition to examining the cluster effect by year of admission, we tested for differences in clustering according to the gender and hospital ward. No differences were noted by gender. However, there were significant differences according to the hospital ward, with medicine having the strongest clustering effect followed by surgery and other wards. Inferences with regard to year were not changed after including the differential clustering according to ward.

**Discussion**

Data presented in this study provide evidence that an important reduction in the average LOS was not associated with a lower level of inappropriate utilisation. As assessed by the AEP, in the hospital studied a shorter LOS was coupled with a higher proportion of inappropriate admissions, and the probability of inactive or inappropriate days of care in 1990 was no lower than that in 1988. The results of the present study support a previous suggestion that cost containment efforts directed at limiting the length of hospital stay may not reduce inappropriate hospital use.13

More patients were admitted to the hospital in 1990 than in 1988, and the proportion of inappropriately admitted patients increased. Also, in 1990 the probability of having at least one inactive day during a patient's stay was
As a measure of goodness of fit of the extended model, we compared the expected and the observed proportions of having inactive days during the hospital stay. The goodness of fit was appropriate. Specifically, the difference between observed and expected proportions was less than 0.8% in all but 3 of 32 comparisons.

As mentioned earlier, the extension of the logistic regression model to allow for random effects made possible the determination and testing of clustering of inactive days. As trend in inappropriateness in our hospital between 1988 and 1990 could be interpreted as the result of specific managerial changes that occurred during that period. For instance, starting in 1989, hospital surgeons were encouraged to operate during the evenings on a fee for service basis. As a result of this, there was an increase in admissions for elective surgery (eg, cataract surgery – data not shown). Some of these admissions could be premature, (thus increasing inappropriateness in admissions), but with few inactive days associated

higher than in 1988, apparently suggesting that the hospital failed to reduce inappropriate stays.

In both years evaluated, days deemed to be inactive followed a trend towards clustering – that is, some patients tended to accumulate inactive days once they had one inactive day during their stay. While the changes that occurred between 1988 and 1990 led to an increase in the probability of having at least one inactive day during a patient’s stay, the within patient clustering of inactive days was lower in 1990. As a consequence, since more patients were inappropriately admitted and those appropriately admitted did not present fewer inactive days, our data indicate that managerial efforts that produced a reduction in the average LOS in 1990 failed to increase hospital efficiency. On the other hand, efforts were partially successful because there were fewer inactive days in a given patient.

An advantage of using data on the number of inactive days in the full LOS in our study is that it allowed for the quantification of clustering of inactive days. This is in contrast to the commonly used approach of selecting one day at random, which is intrinsically limited, as a way of assessing this putative clustering. The analytical approach used in the present paper adequately deals with autocorrelation of the data and adds new information that previous studies which independently analysed admissions and days of stay could not provide. As earlier, the extension of the logistic regression model to allow for random effects made possible the determination and testing of clustering of inactive days.

Trends in inappropriateness in our hospital between 1988 and 1990 could be interpreted as the result of specific managerial changes that occurred during that period. For instance, starting in 1989, hospital surgeons were encouraged to operate during the evenings on a fee for service basis. As a result of this, there was an increase in admissions for elective surgery (eg, cataract surgery – data not shown). Some of these admissions could be premature, (thus increasing inappropriateness in admissions), but with few inactive days associated
Appendix

LIST OF MODIFICATIONS OF THE AEP CRITERIA FOR ITS USE IN THE UNIVERSITARI DEL MAR HOSPITAL

(1) Specifications in the appropriateness criteria

(A9) In the case of insulin treatment, only NPH insulin adjustment was automatically considered a criterion of appropriateness. In the case of slow or semi-slow insulin adjustment, only important changes (more than 5 IU) were considered a criterion of appropriateness.

(B4) Prophylactic subcutaneous injection of heparin as the only treatment was excluded because of routine preventive practice.

(2) List of conditions about which to consult the clinical-epidemiological panel

- **Medicine:**
  - Stroke
  - Enzyme monitoring in myocardial infarction
  - Pancreatitis
  - Tuberculosis plus social problems

- **Surgery:**
  - Colon tumour
  - Gastric ulcer
  - Laparotomy
  - Umbilical hernia
  - Pacemaker implantation

- **Obstetrics/gynaecology**
  - hysterectomy
  - Pelvic inflammatory disease

- **Orthopaedics:**
  - Hip replacement or fracture
  - Knee replacement
  - Femur fracture
  - Laminectomy
  - Disc hernia
  - Hallux valgus

implemented to increase hospital efficiency. Our findings stress the importance of analysing both length of stay and appropriateness when reviewing hospital utilisation.

The AEP Group of the Hospital Universitari del Mar (HUM) is formed by Drs. J. Ballabé (HUM Service of Orthopedics), J. Bañeres (Department Epidemiology and Public Health, IMIM), J. Broquetas (HUM Service of Pneumology), X. Castells (Service of Studies, IMAS), J. Coll (HUM Service of Internal Medicine), M. Sáez (Department of Economics, Pompeu Fabra University, Barcelona), and A. Sigenes (HUM Service of Surgery).

This project was funded by the Institut Municipal d'Assistència Sanitaria (IMAS). A. Muñoz, was partially funded by a grant of the Spanish Ministry of Education and Science (ESAB 93-0150).

The authors thank the physician reviewers (Anna Hernandez, Montse Guizard, Dolores Oller, and Pau Perez) and Dave McFarlane for his help during data analysis. A previous version of this paper was presented at the Fifth European Health Services Research Conference, Maastricht, December 1993.


