Effectiveness of feedback to physicians in reducing inappropriate use of hospitalization: a study in a Spanish hospital

CARLES MOYA-RUIZ1,2, SALVADOR PEIRO´2,3 AND RICARD MENEU2

1Hospital Sant Boi, Barcelona, 2Fundación Instituto de Investigación en Servicios de Salud, Valencia, 3Escuela Valenciana de Estudios para la Salud, Valencia, Spain

Abstract

Objectives. To evaluate the effectiveness of feedback to medical staff in reducing inappropriate hospital days, particularly those attributable to conservative medical discharge policies.

Design. Quasi-experimental pre-test/post-test with non-equivalent control group.

Setting. A publicly funded hospital in industrial belt in Barcelona (Spain), serving a predominantly urban population of 100,000.

Study participants. Two non-equivalent groups: control group (surgery department) and intervention group (internal medicine department).

Intervention. Meetings between hospital management and medical staff of the intervention group to inform clinicians of percentages and reasons for inappropriate stays in their departments.

Main outcome measures. Total inappropriate hospital days and percentage attributable to physicians, measured with the Appropriateness Evaluation Protocol before, during, and after intervention.

Results. There were no relevant differences in the characteristics of the populations whose stays were reviewed during each of the periods. The total number of inappropriate stays and the percentage attributable to the doctor in the control group did not show any differences between the periods. In the intervention group, inappropriate stays attributable to the doctor decreased from 35.9% in the period prior to intervention to 27.7% during intervention (relative drop of 22.8%; P < 0.01), and rose to 32.7% after intervention. Differences in total inappropriate days were not significant.

Conclusions. Providing physicians with feedback about percentages of inappropriate hospital days produced a significant reduction in the number of inappropriate stays attributable to the doctor, although the impact on overall inappropriate stays is inconclusive.

Keywords: feedback, inappropriate hospitalization, utilization review

Inappropriate hospital utilization is defined as the hospitalization of patients who, from the clinical perspective, could be managed on a less intensive health care level [1]. This inappropriate use of resources constitutes an important problem for the health care systems of a large part of the world’s industrialized nations [1–3]. Numerous articles have been published on the subject over the last decade in the US, Canada, and Europe [4–12], and almost 50 papers have addressed this issue in Spain alone [12–14]. Estimates place inappropriate hospitalizations as accounting for one-quarter of admissions, and more than one-third of hospitalization days. In absolute figures, these statistics indicate that >10 million of the somewhat >30 million hospital stays that are registered each year in the Spanish public hospitals [15] are potentially avoidable. Almost all of the Spanish studies report rates of inappropriate stays of >75% due
to delays in discharges once patients no longer require hospitalization.

These findings are spurring researchers (and clinical and health care managers) to search for effective ways to reduce inappropriate hospitalizations imputable to excessively conservative discharge policies. Strikingly, the literature includes very few studies of what is being done to remedy the situation. We could find no studies published in the European Union to date. This situation is all the more surprising in light of the obvious relevance of the problem and of the numerous studies that have revealed its magnitude. Various authors [16, 17] postulate that this may be because in Europe, as opposed to the US, tools for managing hospital utilization are very infrequently used in real health care practice but are, rather, confined to research. The aim of this paper is to evaluate the effectiveness of feedback to physicians about their own percentages of inappropriate hospitalization and the reasons for it, in reducing the rate of inappropriate hospital days.

Materials and methods

The study used a quasi-experimental pre-test/post-test design with a non-equivalent control group. Our research hypothesis was that feedback provided to the physicians through specific meetings would trigger a significant reduction in inappropriate hospital stays, particularly those attributable to conservative medical practices. The study was conducted in a hospital in Barcelona’s industrial belt, which serves a predominantly urban population of 100,000 people. The hospital forms part of Catalonia’s network of public healthcare services, and 95% of its activities are financed with public funds. The center has 130 acute care and 17 convalescence beds that can be occupied by any of the different departments, depending on needs. Although traditionally the hospital was reimbursed on a per diem basis, since 1997 the Catalonian Department of Health has implemented a payment plan that is based, in part, on diagnosis-related groups.

The Internal Medicine Department was the intervention group, and the General Surgery Department of the same hospital, the control group. The intervention consisted of an initial meeting between the hospital’s Medical Director and the doctors participating in the intervention group, in which the concept of inappropriate hospital utilization was explained, along with its causes, implications, and instruments available to measure the phenomenon. The figures on and reasons for inappropriate stays in the pre-test measurement were discussed in this meeting. From that point on, a review of inappropriate utilization in all in-patients in one day was made every 2 weeks over a period of 3 months, and information about these measurements was disclosed to the clinicians in several meetings with the Medical Director, approximately every 2 weeks after each new review. In these meetings, the development, and the main causes of the inappropriate stays rate, were discussed, along with possible ways to reduce it. During the intervention period, medical records were reviewed in the hospital ward. From the first meeting on, it was made clear that changes in the inappropriate stays rate were not going to be accompanied by any type of incentives, economic or otherwise, either positive or negative, nor by any type of penalty. In the control group, inappropriate stays were evaluated on the same days, but the medical staff were not given any information.

The outcomes of the study were the percentage of inappropriate stays and the percentage of inappropriate stays attributable to physicians, identified by reviewing patients’ medical records with the Appropriateness Evaluation Protocol (AEP) [18–20]. The variables examined included an evaluation of the stay (classified as appropriate or inappropriate, and the specification of the cause when a stay was deemed inappropriate), the age and sex of the patient, type of admission (emergency or programmed), length of stay (LOS), and day of the week. Diagnosis was not taken into account because, firstly, the AEP was designed to be used independently of diagnosis and secondly, the size of the sample did not permit an analysis by diagnostic groups.

Sample size was calculated to detect differences from 35% to 25% (one-sided) in the percentage of inappropriate stays between the pre-intervention and intervention period in the Medical ward, with an alpha error of 0.05 and power of 0.80 (280 days in each period). To obtain this sample a review was conducted of all of the stays corresponding to the patients admitted in both wards during 21 days distributed fortnightly between 1 October 1997 and 29 July 1998 (with the exception of Christmas week). Nine days corresponded to the period prior to the intervention, six to the intervention period, and six to the period after the intervention had ceased. Cases whose review corresponded to the day of admission (the AEP uses a different protocol for these days) or the day of discharge were excluded from the sample. A total of 434 stays were reviewed in the control group (pre-intervention: 176; during intervention: 150; post-intervention: 108), while a total of 818 stays were evaluated in the intervention group (pre-intervention: 300; during intervention: 314; post-intervention: 199). The LOS reviewed according to this sample design represents the hospital’s occupation, but is not representative of the LOS in the hospital [21–23].

We used a previously validated Spanish version of the AEP for medical and surgical adult patients [18,19]. The AEP is a well-known independent diagnostic tool, widely used to identify inappropriate hospital stays in non-psychiatric and non-obstetrics adult patients. The AEP was designed to be used by non-medical reviewers, typically trained nurses who have access to medical staff in cases of ambiguity. It should be pointed out that the AEP does not evaluate whether the care provided is appropriate or not, but rather the appropriateness of keeping the patient in the hospital to provide the care. The AEP includes a list of reasons for inappropriate use that is divided into two broad sections: (i) patients who required further hospitalization for medical reasons, but who, on the day their cases were reviewed, received no services that required them to be hospitalized; and (ii) patients who no longer required hospitalization. This division is useful, since the first group suggests that organizational problems at the hospital are keeping people in longer than necessary, while the second group suggests
that there are problems related to the clinical management of discharges, or that the patient is experiencing social or family-related problems that are delaying discharge. The first group includes problems related to scheduling of operating theatres, non-operatory diagnostic procedures, inappropriate pre-operatory stays, problems of availability of operating theatres, procedures that cannot be performed on weekends, delays in receiving test results to determine whether discharge is appropriate, and so on. In the second section three groups of causes have been established, depending on whether the delay in discharge is attributable to the attending physician, the patient himself or his family, or to deficiencies in the healthcare network apart from hospital facilities.

The clinical records were reviewed independently by one doctor and two nurses who had previously received training with records that were not included in the study. This training included the independent review of 41 stays in which the usual concordance indexes were analyzed [24], showing simple agreement levels of 90% or above with the reviewer setting the pattern in the classification of the stays as appropriate or inappropriate, and kappa statistical values that indicated a high level of agreement among the reviewers (Table 1). All of the documents contained in the clinical records were available for review, including medical orders, nursing notes and graphs, inter-consultation notes, diagnostic and therapeutic tests, and laboratory test results. An average of 10 minutes was devoted to reviewing each record. During the pre- and post-test periods, the review was retrospective; during the intervention period, the review was conducted on the basis of records available in the wards. None of the documents or the database used for analytical purposes included any information that made it possible to identify the patients. All the records bore an identification number that permitted retrieving the file for further study whenever necessary.

A descriptive analysis was made of the cases reviewed in both the intervention and the control groups during the three periods, whereas differences between the periods where evaluated with the $\chi^2$ test (substituted in pertinent situations with Fisher’s exact test). The percentage of inappropriate utilization was calculated overall, then broken down in terms of inappropriate use due to the physicians, with a confidence interval of 95% (CI 95%) for the intervention and control groups during each of the periods under study. Differences between the periods were then analyzed. The analysis was based on comparing the results between the pre- and post-intervention periods in each of the groups, and no comparison was made between the groups. The control group was used only for the purposes of monitoring trends, or detecting other concurrent hospital factors influencing results. In line with the direction of our previous research hypothesis, we used one-tailed tests. All the analyses were made using the STATATA™ statistical package (Stata Corporation, University Drive East, College Station, TX, USA).

### Results

Table 2 presents the overall characteristics of the population whose stays were analyzed in both groups during each period. The percentage of patients, broken down by age, sex, and type of admission, showed no significant differences over the three periods. However, as was expected because of the sample design, there were differences in terms of the days of the week examined.

Table 3 shows the overall percentage of inappropriate stays calculated for each group and period. In the control group, the percentage of inappropriate stays reached 30.7% in the pre-intervention period, 29.3% during the intervention phase, and 33.3% after intervention, with no significant differences between the periods. In the intervention group, the percentage of inappropriate stays decreased from 40.7% in the pre-intervention period to 35.3% during intervention, and rose slightly to 37.2% after intervention. These differences also did not reach statistical significance. Table 4 provides a breakdown of inappropriate stays by causes. In both the experimental and the control groups, inappropriate stays attributable to the physician or the hospital’s organizational policies when the patient was ready to go home was the most frequent cause (24–30% of the total number of stays in the controls and 27–30% in the intervention group). Inappropriate stays that were attributable to the family or the patient accounted for 2–3% of all stays in the intervention group and were close to zero in the control group, while the impact of the lack of alternatives to hospitalization was also practically null. Inappropriate stays attributable to scheduling problems were also minimal in both groups during the three periods studied.

Table 5 shows the percentages of inappropriate stays
Table 2 Characteristics of cases reviewed in control and intervention groups in periods before, during, and after intervention

<table>
<thead>
<tr>
<th></th>
<th>Surgery (control group)</th>
<th>Medicine (intervention group)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before intervention</td>
<td>During intervention</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (P ns)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;45 years</td>
<td>46 (26.14)</td>
<td>39 (26.00)</td>
</tr>
<tr>
<td>45–64 years</td>
<td>68 (38.64)</td>
<td>47 (31.33)</td>
</tr>
<tr>
<td>65–79 years</td>
<td>43 (24.43)</td>
<td>41 (27.33)</td>
</tr>
<tr>
<td>&gt;79 years</td>
<td>19 (10.80)</td>
<td>23 (15.33)</td>
</tr>
<tr>
<td>Sex (P ns)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>98 (55.68)</td>
<td>76 (50.67)</td>
</tr>
<tr>
<td>Women</td>
<td>78 (44.32)</td>
<td>74 (49.33)</td>
</tr>
<tr>
<td>Admission (P ns)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programmed</td>
<td>80 (45.45)</td>
<td>68 (45.64)</td>
</tr>
<tr>
<td>Emergency</td>
<td>96 (54.55)</td>
<td>81 (54.36)</td>
</tr>
<tr>
<td>Day (P&lt;0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday</td>
<td>26 (14.77)</td>
<td>29 (19.33)</td>
</tr>
<tr>
<td>Tuesday</td>
<td>0 (0.00)</td>
<td>18 (12.00)</td>
</tr>
<tr>
<td>Wednesday</td>
<td>64 (36.36)</td>
<td>22 (14.67)</td>
</tr>
<tr>
<td>Thursday</td>
<td>48 (27.27)</td>
<td>22 (14.67)</td>
</tr>
<tr>
<td>Friday</td>
<td>0 (0.00)</td>
<td>23 (15.33)</td>
</tr>
<tr>
<td>Saturday</td>
<td>18 (10.23)</td>
<td>17 (11.33)</td>
</tr>
<tr>
<td>Sunday</td>
<td>20 (11.36)</td>
<td>19 (12.67)</td>
</tr>
<tr>
<td>Total</td>
<td>176 (100.0)</td>
<td>150 (100.0)</td>
</tr>
</tbody>
</table>

ns, not significant.

\(P\) corresponds to the \(\chi^2\) test, with the exception of day of the week, when Fisher’s exactness test was used.

Table 3 Percentage of inappropriate days due to all causes in the control and intervention groups in periods before, during, and after intervention

<table>
<thead>
<tr>
<th></th>
<th>Surgery (control group)</th>
<th>Medicine (intervention group)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% CI 95%</td>
<td>% CI 95%</td>
</tr>
<tr>
<td>Prior to intervention</td>
<td>30.68 23.80 37.56</td>
<td>40.66 35.11 46.20</td>
</tr>
<tr>
<td>During intervention</td>
<td>29.33 21.96 36.70</td>
<td>35.35 29.96 40.74</td>
</tr>
<tr>
<td>After intervention</td>
<td>33.33 24.30 42.37</td>
<td>37.19 30.41 43.96</td>
</tr>
<tr>
<td>Difference before-during intervention (P ns)</td>
<td>-1.34 -11.40 8.71</td>
<td>-5.30 -2.40 13.02</td>
</tr>
<tr>
<td>Difference during-after intervention (P ns)</td>
<td>4.00 -7.52 15.53</td>
<td>1.83 -0.67 10.47</td>
</tr>
<tr>
<td>Difference before-after intervention (P ns)</td>
<td>2.65 -8.57 13.88</td>
<td>-3.47 -12.23 5.29</td>
</tr>
<tr>
<td>Total observation period</td>
<td>30.88 26.51 35.24</td>
<td>37.88 34.43 41.12</td>
</tr>
</tbody>
</table>

CI, confidence interval; ns, not significant.

1\(n=434\) (before, \(n=176\); during, \(n=150\); after, \(n=108\)).

2\(n=818\) (before, \(n=305\); during, \(n=314\); after, \(n=199\)).

attributable to either the physician or the hospital, with a breakdown by group and period. In the control group, inappropriate stays amounted to 26.1% in the period prior to intervention, 24.7% during intervention, and 30.6% after the intervention phase, with no significant differences between periods. In the intervention group, the rate of inappropriate use due to the physician or the hospital decreased from 35.9% pre-intervention to 27.7% during intervention (which translates into a relative reduction of 22.8% of inappropriate stays due to these causes), and rose to 32.7% after intervention. The difference was significant when the periods prior to and during intervention were compared (\(P<0.01\), one-tailed test), although the confidence interval (0.9–15.6%) was broad.

**Discussion**

Our hypothesis was that feedback to physicians would reduce inappropriate hospital stays attributable to conservative medical practices. We believed that it was unlikely that the feedback which was given to the medical staff would have
any impact on other problems such as scheduling, those basically related to diagnostic services, issues related to the family, or the lack of health care facilities other than hospitals. Studies conducted in Spain indicate that >75% of inappropriate stays are due to medical decisions and we therefore assumed that a reduction in stays of this nature would produce a reduction in inappropriate stays overall. Our results show that feedback to medical staff did produce an absolute reduction of 8.5% of inappropriate stays attributable to physicians (from 35.9% to 27.4%), but the impact on the overall rate of inappropriate stays was only 5.4% (from 40.7% to 35.3%), and was not statistically significant. It should be pointed out that even a reduction of as little as 5% of inappropriate stays can have an important economic impact for hospitals. However, we have no power to detect the statistical significance of this reduction (∼1200 days in each
period); therefore, our results are inconclusive in terms of this. A possible explanation for the minor impact detected on overall inappropriate days could be that more thorough annotation in the clinical records of family and social problems – a probable occurrence as part of the Hawthorne effect usually associated with the knowledge that an evaluation is taking place – may have produced a shift in the assignment of the causes from the physicians to the family or issues related to the alternative health care network. It is likely that this phenomenon did occur, since in the intervention group inappropriate stays due to family-related problems increased from 1.9% to 3.2%, and stays due to problems with the provision of alternative care from 0.6% to 1%. However, these percentages do not carry sufficient weight to explain the differences detected in the reduction of inappropriate stays attributable to the medical staff when contrasted with the overall reduction. Another explanation might be that the group responded to the feedback provided, intentionally or otherwise, with opportunistic behaviour, and increased the therapeutic intensity of aspects that implied changes in the evaluation of the stays. This possibility derives from the sensitivity of AEP to therapeutic intensity, and the fact that the doctors were made privy to the way the AEP works as part of the intervention. If this phenomenon indeed occurred, it would have been accompanied by the paradox of an apparent reduction in the number of inappropriate stays, together with an overuse of therapeutic processes, and a corresponding rise in costs and loss of quality. However, an assessment of expenditures in pharmaceutics and consumable materials during the different periods studied revealed no relevant variations. It is most likely that a combination of factors, including those outlined above, contributed to these inconclusive results with respect to the reduction of inappropriate stays overall.

In terms of the possible impact of the reduction of inappropriate stays on the average LOS in the different departments, the LOS showed a downward trend in the intervention group after intervention (from 8 days in October to 7 days during the intervention period and 6 days in the post-intervention period, according to the hospital’s indicator system), while no changes were seen in the control group. Although these data may suggest that the intervention did have an impact on the LOS, the design of the study does not make it possible to make a conclusive evaluation of this point. The few studies that have examined the impact of feedback on reducing the incidence of inappropriate hospital utilization have shown results that were clearly positive [25–29]. This is partly because, in these studies, measures were associated with economic incentives or more intensive interventions. For example, in the study conducted by Mozes et al. [27], physicians were required to review their patients with the AEP on a daily basis and provide written justification for cases where the patient remained hospitalized without meeting any of the AEP criteria. A predictable bias in favor of publishing positive results may also have played a role [30]. Analysis of the reasons behind inappropriate stays reveals that, in line with other studies [31,32], the low number of inappropriate days imputable to the patient or the family contradicts the extended belief that these factors are responsible for a significant proportion of inappropriate hospital occupation. Even when taking into account the percentage of inappropriate stays that are deemed to be due to deficiencies in the health care network in general (such as a lack of alternative health care facilities), >90% of the inappropriate stays exposed in our study were due either to the physician or to the hospital.

When interpreting results, it is important to take into consideration some limitations. Foremost among these is the fact that the groups are not comparable. Although this study analyzed the same groups during different periods, and did not compare one group with the other, our results may have been sensitive to possible differential influences. A second limitation that one would expect is the possible contamination of the controls through contact with the intervention group, although in our case this seems not to have occurred. Thirdly, while the two nurses who reviewed the majority of the medical records were not aware of the study’s hypothesis, the doctor who reviewed the remainder was one of the researchers and was, therefore, not blinded to the hypothesis; as a Medical Director, this person also participated in the intervention. Given the characteristics of the AEP (it uses explicit criteria and has high reliability), it is very unlikely that this aspect will have influenced the results. The fact that there was a high degree of concordance between the two nurses and the researcher during training prior to the study further indicates that this factor had little effect on the results. Other limitations are related to the foreseeable sensitivity of the sample design to weekly and seasonal variations.

This is not the place to comment on the limitations generic to the AEP and utilization reviews; these are described extensively elsewhere [1,3,11,33,34]. These limitations, nonetheless, do not imply that identifying and reducing the inappropriate utilization of hospital care is not a worthy endeavour (as long as the necessary quality of care can be maintained). This exercise can help reduce the unit costs of hospital processes, as well as improve the quality of the care provided, and prevent the onset of problems associated with prolonged hospitalization (increase in risk of nosocomial infections, ‘hospitalism’, loss of quality of life and inconvenience for patients, and expenditures on the part of the family) and can help to shorten diagnostic periods and eliminate delays in treatment. The findings, echoed in several studies, that one out of every three hospital stays is clinically inappropriate is not a fact that can be ignored, as it affects physicians, healthcare managers, and policy makers alike. On the contrary, these figures suggest that there is ample room for improving efficiency while maintaining quality care.

Although it is no doubt true that solving this problem will require working on many fronts, it is also clearly possible to tackle the problem on the local level with different types of interventions: changing procedures for requesting and processing diagnostic tests, updating hospital organization systems in order to be able to perform diagnostic tests under an out-patient regime within a reasonable time frame,
correcting scheduling of operating theatres, conducting pre-
surgery evaluations under an outpatient regime, extending
the use of surgery procedures that do not require hos-
pitalization, and exploiting ‘hospital at home’ arrangements
(whatever better co-ordination with primary care teams).
While it will be necessary in many cases to balance the costs
of these interventions with the benefits derived from them,
in others these measures are simply part of adequate medical
management schemes. This study has found that providing
feedback to physicians is effective in reducing the rate of
inappropriate stays attributable to medical staff, and suggests
that intervention in the hospital at large or, in a given
department, can have a positive impact, precisely because it
places decision-making responsibility in the hands of the
medical staff and hospital managers. Our results highlight
that this step is crucial if the problem of inappropriate
utilization of hospital resources is to transcend the pages of
research publications and enter the realm of day-to-day
hospital management and the organization of medical and
surgical departments.

Acknowledgements

Our gratitude to Montserrat Rios Fernandez and Josefa Diaz-
Flores, who reviewed a large part of the clinical records.

References

1. Payne SMC. Identifying and managing inappropriate hospital

2. Liberati A, Apolone G, Lang T, Lorenzo S. A European
project assessing the appropriateness of hospital utilization:
Background, objectives and preliminary results. Int J Qual Health

3. Peiró S, Portella E. Identification of the inappropriate use of
hospitalization: the search for efficiency (in Spanish). Med Clin

4. Booth BM, Ludke RL, Wakefield DS, Kern DC, Fisher EM,
Ford TW. Nonacute inpatient admission to Department of
Veterans Affairs Medical Centers. Med Care 1991; 29 (suppl.):
AS40–AS50.

5. Booth BM, Ludke RL, Wakefield DS et al. Nonacute days of
care within Department of Veterans Affairs Medical Centers.
Med Care 1991; 29 (suppl.): AS51–AS63.

6. Smith CB, Goldman RL, Martin DC et al. Overutilization of
acute-care beds in Veterans Affairs Hospitals. Med Care 1996;
34: 85–96.

7. Lang T, Davido A, Logerot H, Meyers L. Appropriateness of
7: 233–238.

hospital use: an overview of Italian studies. Int J Qual Health

9. Bentes M, Gonsalves ML, Santos M, Pina E. Design and
development of a utilization review program in Portugal. Int J

10. Santos-Eggiman B, Paccaud F, Blane T. Medical appropriateness
of hospital utilization: an overview of the Swiss experience. Int

S. The hospital admission study in England: are there alternatives
to emergency hospital admission? J Epidemiol Comm Health 1995;
49: 194–199.

12. Lorenzo S, Suñol R. An overview of Spanish studies on ap-
propriateness of hospital use. Int J Qual Health Care 1995; 7:
213–218.

13. Lorenzo S. Resources utilization review. Spanish studies (in

14. Meneu R, Peiró S. Inappropriate hospital utilization review in
Spain: from scientific communication to practical use? (in

15. Instituto Nacional de Estadística. Survey of Health Centers With

16. Restuccia JD. The evolution of hospital utilization review
methods in the United States. Int J Qual Health Care 1995; 7:
253–260.

17. Meneu R. Does reduction of inappropriate hospital use reduce
health expenditures? What do we know about it. International
Conference on The Evaluation of Appropriateness of Hospital Use:
Its Relevance for Controlling the Delivery and Financing of the Hospital
Care in Europe. 20 March 1997, Florence, Italy.

18. Gertman PM, Restuccia JD. The Appropriateness Evaluation
Protocol: a technique for assessing inappropriate days of hospital


20. Peiró S, Meneu R, Rosello ML et al. Validity of the protocol
for evaluating the inappropriate use of hospitalization (in Span-


22. Gange SJ, Muñoz A, Saez M, Alonso J. Use of the beta-
binomial distribution to model the effect of policy changes on

23. Peiró S, Meneu R. Inappropriate use of hospitalization: problems

24. Fleiss JL. The measurement of interrater agreement. In: Fleiss
JL, ed. Statistical Methods for Rates and Proportions. New York,

25. Restuccia JD. The effect of concurrent feedback in reducing

GO. Improving physician performance through peer compar-


Accepted for publication 2 April 2002