Assessing the necessity of hospital stay by means of the Appropriateness Evaluation Protocol: a different perspective

To the Editor: Smeets et al. [1] state that the aim of their article recently published in the *International Journal for Quality in Health Care* was ‘to review the Appropriateness Evaluation Protocol (AEP) with respect to assessing the necessity of hospital stay at the University Hospital of Maastricht’. However, the study is in fact a narrative review of some of the clinimetric properties of the AEP (essentially reliability and criterion validity), followed by a subjective interpretation of the impact of these properties on the practical use of the instrument, which leads the authors to ‘not encourage further use until additional improvements to both the instrument and the review conditions have been made’.

In our opinion, not only does the study contain various inaccuracies and inconsistencies, the framework it establishes is not suitable for the evaluation of the practical implications of the AEP’s clinimetric properties, a circumstance that has led authors to reach conclusions that are, at the very least, debatable.

Some of the inaccuracies we would like to point out affect both the selection and interpretation of the studies. These are discussed below. In the section outlining the search of the literature, the text states:

> Of the remaining studies addressing the appropriateness of stay measured by the AEP, five were in Spanish thus limiting the accessibility of the results. Two of these concerned surgical departments but did not mention whether they used the Elective Surgery version of the AEP. (p. 342)

One of these two references corresponds to an article drafted by several of the authors of this letter [2], which was published in English in the *International Journal for Quality in Health Care*, and contains a description of the version of the AEP used in the ‘Materials and methods’ section. What is more, Smeets et al’s oversight is doubly surprising because our study, as the title itself reflects, consisted of a review of appendectomies, where it would have been very difficult to use the Elective Surgery version of the AEP.

By using the search criterion ‘Appropriateness Evaluation Protocol’ on Medline on 13 March 2001, we found 129 articles; 59 more than the 70 encountered by Smeets et al. In spite of basing their search on broader terms, these authors appear to have employed a search strategy that failed to detect numerous relevant studies, including works that validated the AEP in different European contexts [3–5], such as the reliability study of the European version [6].

The conclusion that ‘feedback to physicians can lead to a small decrease in unnecessary days’ is accompanied by nine bibliographical references, of which at least seven are studies that neither evaluated nor used any kind of feedback at all. The authors, however, omitted relevant publications in this area [7–9], including studies that evaluated the impact of utilization review on the management of health care centres [10]. Another uncited study concluded that retrospective feedback reduced inappropriate admissions by over 50% compared with the level before feedback [11]. Smeets et al. state that ‘retrospective feedback can reduce unnecessary stays by at most 10%’. A closer reading of this controlled trial, however, indicates that retrospective feedback resulted in nearly a 12% reduction in unnecessary stays and, when ‘effectively’ delivered (i.e. delivered to attending physicians), over 25% reduction [12].

The authors do not indicate the significance of ‘a small decrease’, another important aspect, since a small decrease from, for example, 30 to 28% of inappropriate stays would translate into preventing more than 60 000 unnecessary stays yearly in the context of the Spanish public health care system.

The inclusion in Table 6 [1] of Querido’s study, a review conducted with implicit methods in 20 Dutch hospitals and published 18 years before the original publication of the AEP [13], is a further inaccuracy that requires correction.

These points notwithstanding, perhaps the greatest problem resides in the fact that the authors go from recognizing the ‘comparatively good performance’ of the AEP reported in the literature, to concluding that its validity and reliability are ‘not yet up to standard’ and recommending that it not be used. Nonetheless, the AEP, or similar instruments, are used in many ways, and results can be affected differently by the instrument’s clinimetric properties, which can be summarized as follows: (1) high reliability between reviewers, (2) moderate criterion validity, with a tendency to over-estimate inappropriate use of hospitalization compared with the implicit judgement of clinicians, (3) a certain susceptibility to different sample designs and conditions under which the review is conducted.

The AEP, and similar instruments that have been studied
to a lesser extent, have been used both by hospital care providers and financiers. Financiers and regulators have used this type of instrument to:

1. control the supply of new hospital services (deciding whether to grant a Certificate of Need) when there is a significant volume of idle resources;
2. plan the supply of alternative services (i.e. build a community health care centre rather than expand a hospital when saturation of the hospital is due to inappropriate stays corresponding to patients who could be served in another setting);
3. adjust reimbursement, usually by adjusting payment per stay according to the overall percentage of inappropriate use and, on other occasions, refusing to reimburse specific stays deemed to have been inappropriate;
4. develop hospital profiling to select providers; and
5. develop standards for the average ‘appropriate’ stay associated with different processes in order to have criteria which can be used to compare with the real average stay associated with each process.

Healthcare providers have used these instruments to:

1. identify problems related to the hospital’s organization or management of patients that, if solved, would reduce the percentage of inappropriate stays (i.e. a reduction in pre-operative stays for elective surgery; diagnosis of certain cases on an outpatient basis; incorporation of planned discharge programs);
2. monitor percentages of inappropriate use over time and, in comparison with other hospitals, to detect variations and learn from those with the best results;
3. improve quality of service and avoid unnecessary prolongation of stays and their corresponding risks (e.g. nosocomial infections), discomfort and expenditure on the part of the patient, and improve scheduling and timing of diagnostic procedures;
4. identify candidates for discharge and to propose medical evaluation and/or schedule discharge planning for them.

In general, all uses based on the percentages of unnecessary stays would barely be affected by problems related to the instrument’s moderate validity and biases associated with the review methodology discussed in this study. Uses based on the individual evaluation of a specific patient, such as to deny reimbursement of a specific stay or to identify preliminary candidates for discharge, would be more debatable in terms of the clinimetric properties of the AEP. Put another way, if a review indicates that a hospital has a rate of inappropriate admissions for elective surgery of 35%, of which one-third are due, for example, to inappropriate handling of admission for selective surgery, this hospital would do well to implement measures to correct the problem, even if the ‘real’ percentage of inappropriate stays is 30 or 25%; on the other hand, no hospital should discharge a patient solely on the basis of a review with the AEP.

As has already been stated [14], the main difference between utilization reviews in the USA and Europe is that while in the USA the instrument is used for hospital management, in Europe its use is limited to research for the exclusive purpose of scientific study. However, and after more than 12 years of published European research on all manner of aspects related to the AEP, it is considered a mature instrument, ready for implementation in routine health care management, and in Spain in fact it has been used for this purpose for the last few years [15]. This is because the usefulness of an instrument is not so much a function of its perfection but of knowledge of how to use it and how to account for its limitations in different contexts.

Salvador Peiró
Escuela Valenciana de Estudios para la Salud and Fundación Instituto de Investigación en Servicios de Salud Valencia, Spain
Ricardo Meneu
Fundación Instituto de Investigación en Servicios de Salud Valencia, Spain
Susana Lorenzo
Fundación Hospital Alcorcón Spain
Joseph D. Restuccia
Boston University School of Management Boston, MA, USA

References


