

Who should be admitted to hospital? Evaluation of a screening tool

LP Leung 梁令邦
KL Fan 樊潔玲

Objective	To develop a tool for evaluating the appropriateness of acute hospital admissions in Hong Kong and test its reliability.
Design	The tool was based on the Appropriateness Evaluation Protocol and consensus of local Emergency Medicine specialists. Reliability was tested through retrospective chart review.
Setting	Tertiary teaching hospital, Hong Kong.
Patients	Seventy-five randomly selected patients, who were admitted to the specialty of Internal Medicine or General Surgery via the Accident and Emergency Department in 2006, were reviewed.
Main outcome measures	The intra-rater and inter-rater agreement on appropriateness of an admission.
Results	A 19-criterion protocol for assessing the appropriateness of acute hospitalisations was constructed. The kappa coefficient for intra-rater agreement was 0.73 (95% confidence interval, 0.58-0.88) and that for inter-rater agreement was 0.67 (95% confidence interval, 0.51-0.83).
Conclusion	The new protocol was shown to have substantial reliability for evaluating whether an acute hospital admission was appropriate. The findings in this study provide a basis for testing the validity of the new protocol as well as determining the extent of inappropriate acute hospital admissions in Hong Kong.

Introduction

The cost of health care delivery has come under scrutiny in recent years. With a relatively fixed health care budget and escalating costs of health service provision, reducing inappropriate hospital admissions is an appealing approach to addressing the fiscal realities of policy-makers. Determining the appropriateness of hospital admission, however, is problematic, because a gold standard does not exist.

In the literature, there are two distinct methods for evaluating appropriateness. The first uses subjective, usually expert, opinion. The second employs comparatively objective utilisation review tools. For the former, the assessment is critically dependent on the reviewer's judgement. However, reviewers differ in terms of background and training, scope of expertise, and style of practice. This accounts for low inter-rater agreement and discrepancies when it comes to evaluation of the appropriateness of health care.^{1,2} As for utilisation review tools, they were initially developed in the United States. The tools most commonly employed are the Intensity of service, Severity of illness, Discharge screen and Appropriateness review system (ISD-A), and the Appropriateness Evaluation Protocol (AEP). The ISD-A was developed by InterQual, Inc in 1978. The AEP was based on the work of Goldberg and Holloway, with subsequent revisions and updates.³⁻⁵ Both the ISD-A and the AEP use sets of explicit criteria to determine whether an admission is appropriate and whether any given in-patient day is actually required.

There are three local studies on appropriateness of acute admissions.⁶⁻⁸ In two of them, the AEP was adopted without modification. 'Discharge within 24 hours' was used as the criterion for inappropriateness in the other. The rate of inappropriate admissions reported ranged from 4.7% to 10.7%. There are several deficiencies in these studies. First, the performance of AEP in the local setting has never been documented. Second, the AEP was developed in the United States. There are obvious differences in case mix and medical practice between Hong Kong and the United States, and the original version of AEP may not be applicable locally. Third, using 'discharge within 24 hours' as a criterion is far from

Key words
Health services misuse; Hospitalization;
Patient admission; Patient selection;
Reproducibility of results

Hong Kong Med J 2008;14:273-7

Accident and Emergency Department,
Queen Mary Hospital, Pokfulam, Hong
Kong

LP Leung, FHKAM (Emergency Medicine)
KL Fan, FHKAM (Emergency Medicine)

Correspondence to: Dr LP Leung
E-mail: lpleung@yahoo.co.uk

誰適合急症入院？對一套篩選工具的評估

目的	發展一套適用於香港的急症入院適當性評估系統及測試其可靠性。
設計	此工具主要根據《住院適當性評估系統》及急症科專科醫生的共識，並進行回顧性研究以測試其可靠性。
安排	香港一所大學教學醫院。
患者	回顧2006年經急症室轉介至內科或外科的75位隨機抽樣的病人紀錄。
主要結果測量	按入院適當性評分者本身及評分者之間的信度。
結果	本研究建立了一套有19項準則，用作評估急症入院適當性的系統。評分者本身信度的kappa系數為0.73（95%置信區間：0.58-0.88）；評分者之間信度的kappa系數為0.67（95%置信區間：0.51-0.83）。
結論	利用此新系統來評估急症入院的適當性有高度可靠性。此研究為測試新系統的有效性提供了基礎，有助了解香港不適當急症入院的程度。

satisfactory, as there are many factors affecting the length of stay in hospital.

This study was part of a project to develop a tool for evaluating the appropriateness of acute hospital admissions and assessing the prevalence of inappropriate admissions in Hong Kong. The primary objective of this study was to test the reliability of a local version of the AEP for evaluating the appropriateness of acute hospitalisation.

Methods

Tool development

The AEP was chosen as the prototype. This is because it has been extensively tested for reliability and validity. The results were satisfactory with regard to its application in different specialties and in different countries.⁹ A group of four specialists in Emergency Medicine from three different hospitals participated in the tool development process in early 2007. Two of them were consultants, while the remaining two were a senior medical officer and a medical officer. All had over 10 years of experience in the practice of Emergency Medicine. They were asked to add, delete, or modify the 16 criteria of the admission part of the original AEP. The local version (HK-AEP) was drafted once consensus was reached.

Study design

This study was a retrospective chart review of the medical records of a random sample of admitted patients by one reviewer on two occasions 3 months apart (for intra-rater reliability), and two independent

reviewers (for inter-rater reliability). They all used the HK-AEP.

Patients admitted to the specialty of Internal Medicine and General Surgery of a tertiary teaching hospital in 2006, were recruited by proportional random sampling (Internal Medicine: General Surgery=2:1) using the random number table. The ratio corresponded to that of the number of admissions to the respective specialties in 2006. Patients were included if they were 18 years or older, and admitted for reasons other than trauma. Those younger than 18 years, admitted via sources other than the emergency department, non-entitled patients, or trauma patients were excluded.

Reviewers with a specialist qualification in Emergency Medicine and at least 10 years of experience of working in Hong Kong were recruited, and it was a requirement that they should not have had direct involvement in the care of the selected patients.

The medical records of the selected patients were presented in a standardised abstract format. All abstracts were prepared by one of the investigators. The clinical information of the index admission in the abstract was based on the record of the Accident and Emergency Department, supplemented by the discharge summary retrieved from the hospital computer system. Each reviewer was blinded to the other's judgement, as well as the identity and outcome of the admitted patients. All three reviewers assessed the same set of abstracts. Data on the demographic characteristics of the patients, and the number of appropriate and inappropriate admissions as judged by each reviewer (using the HK-AEP) were collected.

Statistics

Descriptive statistics were used to describe the demographic characteristics of patients. Reliability, as reflected by the intra- and inter-rater agreement, was represented by the kappa statistic. A kappa value of at least 0.4 was considered the minimum requirement to support the reliability of the tool. This is because it is the cut-off point for moderate or better agreement in the two most commonly used classification systems of the kappa coefficient.^{10,11} Furthermore, most international studies on the reliability of AEP (original or adapted version) ranged from 0.3 to 0.8.¹²⁻¹⁷ For the determination of inter-rater agreement, two reviewers were chosen. According to Shoukri,¹⁸ when seeking to detect a kappa of 0.4 or greater on a dichotomous variable, it is not advantageous to use more than three reviewers per subject. It was shown that for a fixed number of observations, increasing the number of reviewers beyond three adds little to the power of the hypothesis test or the width of confidence interval. As for the sample

TABLE 1. Patient characteristics

Characteristic	Data
Age (years)	
Mean	67
Range	22-96
Sex	
Male	34 (45%)
Female	41 (55%)
Admission specialty	
Internal Medicine	50 (67%)
General Surgery	25 (33%)
Triage category	
1	0
2	3 (4%)
3	48 (64%)
4	24 (32%)
5	0
Activity of daily living status	
Not specified	3 (4%)
Independent	46 (61%)
Dependent	9 (12%)
Partially dependent	17 (23%)

TABLE 2. Intra-rater agreement*

	First review		Total
	Appropriate	Inappropriate	
Second review			
Appropriate	34	7	41
Inappropriate	3	31	34
Total	37	38	75

* Observed percentage agreement=87%; kappa coefficient (agreement corrected for that expected by chance)=0.73; 95% confidence interval, 0.58-0.88

size of patients,¹⁹ in a two-observer study to detect a kappa of 0.4 or greater with $P \leq 0.05$ as significant, and power 90% in a one-tailed test, the minimum number of subjects required is 54. Statistical analysis was performed using the Statistical Package for the Social Sciences (Windows version 10.0; SPSS Inc, Chicago [IL], United States).

Results

The evaluation tool: Hong Kong version of the Appropriateness Evaluation Protocol

The general structure of the original AEP and its independence from a specific diagnosis were retained in the new HK-AEP (Appendix). The main difference of the HK-AEP was the inclusion of the Emergency Medicine ward (EMW) as a factor for consideration. There were 19 criteria in total. The decision rule was simple; if one of the 19 criteria was met, the admission was considered appropriate.

Patient characteristics

Seventy-five patients were recruited; 50 were

TABLE 3. Inter-rater agreement*

	Reviewer 1		Total
	Appropriate	Inappropriate	
Reviewer 2			
Appropriate	40	1	41
Inappropriate	11	23	34
Total	51	24	75

* Observed percentage agreement=84%; kappa coefficient (agreement corrected for that expected by chance)=0.67; 95% confidence interval, 0.51-0.83

admitted to the specialty of Internal Medicine and 25 to General Surgery. The mean age was 67 years with a slight female predominance. Most of them belonged to triage category 3 (64%) and had independent activity of daily living status (61%) [Table 1].

Reliability

The kappa coefficient for intra-rater agreement was 0.73 (95% confidence interval, 0.58-0.88) [Table 2]. The kappa coefficient for inter-rater agreement was 0.67 (95% confidence interval 0.51-0.83) [Table 3]. Using the Landis and Koch's guidelines for interpreting the kappa value,¹⁰ both intra-rater and inter-rater agreements were regarded as substantial.

Discussion

Determining the appropriateness of an acute admission is a complex process. There is not a single set of criteria that is universally applicable. In spite of this inherent limitation, utilisation review tools like AEP attempt to overcome the vagaries of subjective evaluation by providing a more objective and explicit means of assessment. During the construction of the HK-AEP, an admission was considered inappropriate if given a set of clinical features or diagnostic test results, there was a potentially lower technology alternative to admission to an in-patient bed and the alternative did not impose more harm to the patient. Cost was not considered in this definition.

Because of the difference in health care delivery between the United States and Hong Kong, the original AEP had to be modified before being applied to the local setting. The chief modification made in the HK-AEP was the addition of the EMW factor. The EMW evolved from observation wards of accident and emergency departments. The EMW is a specialty ward manned by Emergency Physicians. For the HK-AEP, four principles of patient selection for EMW admission were adhered to. These were based on recommendations of the American College of Emergency Physicians.²⁰ First, there should be a clear goal for EMW admission, such as to evaluate a symptom that bears a high risk but low probability of a serious outcome. Patients with atypical chest

pain are typical examples. Second, the intensity of service needs had to be limited; the demand for medical and nursing care had to be compatible with the manpower and facilities of an EMW. Third, the severity of illness had to be limited; preferably the patient had to manifest only one disease or one body system affected. Finally, the clinical condition needed to be suitable for observation and conservative treatment.

The HK-AEP was drafted from the perspectives of Emergency Physicians, and based on two considerations. In Hong Kong, the Emergency Physicians are often the sole decision-makers of whether a patient needs admission. Second, the protocol was designed to be independent of any specific clinical condition. An Emergency Physician, being an all-round medical practitioner, was in an advantageous position to determine whether admission was actually required. This is because patients often present with an undifferentiated symptom, not a specific disease. However, with the advent of medical technologies and administrative changes, input from other specialties can be helpful.

Overall, the reliability of the HK-AEP was substantial, and comparable to what was noted in similar studies from elsewhere.¹²⁻¹⁷ In this study, the intra-rater was higher than inter-rater reliability; recall bias of the reviewer in the intra-rater assessment may account for the difference. A longer duration between the two evaluations may have been preferable. Apart from methodological considerations, interpretation of the kappa coefficient was also influenced by the prevalence of the condition studied.²¹ If the prevalence of the condition being studied was high, that is, the sample was more homogenous, the kappa may actually be low. It is recommended that the prevalence of the condition to be rated should not be higher than 50%. Based on the results of the

three local studies mentioned earlier, the rate of inappropriate admission ranged from 5% to 11%. Thus, it is believed that the effect of prevalence on the present study was not significant. Another point of concern was the reliability of the protocol when used in patients in a different setting, eg a different hospital. This concern was reflected in some studies, where the same evaluation tool could not be transferred to other health care settings or even to different hospitals in the same geographical area.²² In Hong Kong, however, all public hospitals operate within the same health care system, which is both equitable and accessible. Thus, the case mix in different hospitals was unlikely to be significantly different. As a result, the reliability of this tool was likely to be reproducible in other local hospitals.

Conclusion

The construction of the HK-AEP was the first step in evaluating the appropriateness of acute hospital admissions in Hong Kong, and has shown substantial reliability. This finding provides a basis for testing its validity in a later study and assessing the prevalence of inappropriate acute admissions. It is hoped that a reliable and valid evaluation tool suitable for local use can then be produced.

Acknowledgements

The authors wish to thank Dr F Ng, Consultant, Accident and Emergency Department (A&E), Caritas Medical Centre; Dr HF Ho, Chief of Service, A&E, Queen Elizabeth Hospital; and Dr SD Yeung, Consultant, A&E, Alice Ho Miu Ling Nethersole Hospital for their invaluable advice and participation in this study. The authors have received a research grant from the Hong Kong College of Emergency Medicine.

References

- Richardson FM. Peer review of medical care. *Med Care* 1972;10:29-39.
- Sanazaro PJ. Quality assessment and quality assurance in medical care. *Annu Rev Public Health* 1980;1:37-68.
- Goldberg GA, Holloway DC. Emphasizing "level of care" over "length of stay" in hospital utilization review. *Med Care* 1975;13:474-85.
- Restuccia JD, Holloway DC. Barriers to appropriate utilization of an acute facility. *Med Care* 1976;14:559-73.
- Gertman PM, Restuccia JD. The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981;19:855-71.
- Lo CM, Yau HH. Appropriateness of emergency admissions in a regional hospital. *Hong Kong J Emerg Med* 1995;2:15-8.
- Li KM, Kwan M. An evaluation study of emergency admission of the United Christian Hospital. *Hong Kong J Emerg Med* 1995;2:10-4.
- Lo CM, Leung SH, Lam CS, Yau HH. Clinical audit on short stay emergency medical admission. *Hong Kong J Emerg Med* 2003;10:30-6.
- Payne SM. Identifying and managing inappropriate hospital utilization: a policy synthesis. *Health Serv Res* 1987;22:709-69.
- Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. NY: John Wiley; 1981.
- Lorenzo S, Lang T, Pastor R, et al. Reliability study of the European appropriateness evaluation protocol. *Int J Qual Health Care* 1999;11:419-24.
- Apolone G, Fellin G, Tampieri A et al. Appropriateness of hospital use: report from an Italy study. *Eur J Public Health* 1997;7:34-9.
- Kaya S, Erdem Y, Dogrusoz S, Halici N. Reliability of a

- hospital utilization review method in Turkey. *Int J Qual Health Care* 1998;10:53-8.
15. Sangha D, Schneeweiss S, Wildner M, et al. Metric properties of the appropriateness evaluation protocol and predictors of inappropriate hospital use in Germany: an approach using longitudinal patient data. *Int J Qual Health Care* 2002;14:483-92.
 16. Bentes M, Gonslaves ML, Sauros M, Pina E. Design and development of a utilization review program in Portugal. *Int J Qual Health Care* 1995;7:201-12.
 17. Santos-Eggimann B, Paccaud F, Blanc T. Medical appropriateness of hospital utilization: an overview of the Swiss experience. *Int J Qual Health Care* 1995;7:227-32.
 18. Shoukri MM. Measure of interobserver agreement. Boca Raton, Fla: Chapman and Hall/CRC; 2004.
 19. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic: confidence interval construction, significance-testing and sample size estimation. *Stat Med* 1992;11:1511-9.
 20. Brillman J, Mathers-Dunbar L, Graff L, et al. Management of observation units. *American College of Emergency Physicians. Ann Emerg Med* 1995;25:823-30.
 21. Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol* 2000;53:499-503.
 22. O'Neill D, Pearson M. Appropriateness of hospital use in the United Kingdom: a review of activity in the field. *Int J Qual Health Care* 1995;7:239-44.

Appendix. Hong Kong version of the Appropriateness Evaluation Protocol

Appropriateness of admission criteria

Severity of illness criteria

1. Sudden onset of unconsciousness or disorientation (coma or unresponsiveness)
2. Pulse rate <50 or >140 per minute
3. Blood pressure (with related symptomatology/complications)
 - systolic <90 mm Hg or >200 mm Hg
 - diastolic <60 mm Hg or >120 mm Hg
4. Acute threat to or loss of sight
5. Acute loss of hearing
6. Acute loss of ability to move body part
7. Acute ataxia
8. Persistent fever
 - 37.8°C (aural or oral) or
 - 38.3°C rectally
 - for >5 days or with positive TOCC* history
9. Acute bleeding
10. Severe electrolyte or blood gas abnormality (any of the following, in mmol/L):
 - sodium <123 or >156
 - potassium <2.5 or >6.0
 - venous bicarbonate (unless chronically abnormal) <20 or >36
 - arterial pH <7.30 or >7.45
11. Electrocardiographic evidence of acute ischaemia
12. Refractory hypoxemia with SpO₂ <90% (unless chronically abnormal)
13. Wound dehiscence or evisceration

Medical procedure

1. Intravenous medications and/or fluid replacement (does not include tube feeding) which cannot be managed in an Emergency Medicine ward
2. Surgery or procedure scheduled within 24 hours requiring
 - general or regional anaesthesia, or
 - use of equipment, facilities, or procedures available only in a ward
3. Vital sign monitoring every 2 hours or more often (may include telemetry or bedside cardiac monitor) which is beyond the scope of management of an Emergency Medicine ward
4. Chemotherapeutic agents that require continuous observation for life-threatening toxic reaction
5. Parenteral antibiotics at least every 8 hours which cannot be managed in an Emergency Medicine ward with a view to switch to oral antibiotics within 24-48 hours
6. Intermittent or continuous ventilatory support (invasive or non-invasive) at least every 8 hours

* TOCC denotes travel, occupation, cluster of individuals with similar symptoms, contact history