

Developing of the Appropriateness Evaluation Protocol for Public Hospitals in Iran

Anvar Esmaili¹; Hamid Ravaghi^{2,*}; Hesam Seyedin³; Bahram Delgoshaei²; Masoud Salehi³

¹Department of Health Management and Economics, School of Public Health, Tehran University of Medical Sciences, Tehran, IR Iran

²Department of Health Services Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, IR Iran

³Health Management and Economics Research Centre, Iran University of Medical Sciences, Tehran, IR Iran

*Corresponding Author: Hamid Ravaghi, Department of Health Service Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, IR Iran. Tel: +98-2188793805, E-mail: ravaghi.h@iums.ac.ir

Received: March 17, 2014; Revised: April 28, 2014; Accepted: July 2, 2014

Background: Employment of utilization review instruments is a method for managing costs and efficiency in the healthcare systems. **Objectives:** This study developed an instrument for measuring the level of inappropriate acute hospital admissions and days of care in Iran. **Patients and Methods:** The American version of the Appropriateness Evaluation Protocol (AEP) was modified, using the agreement method, by a multidisciplinary group of physicians. We conducted a retrospective descriptive study of 273 randomly selected patients admitted to Imam Khomeini Hospital of Tehran University of Medical Sciences in Tehran, Iran. For the reliability study, two nurses were asked to review patients' medical records using the instrument. Validity was appraised by pairs of clinicians, including two general surgeons, two internists and two gynecologists. The degree of consensus between the three pairs of clinicians was compared with that of the nurses. **Results:** Inter-rater and intra-rater reliability testing revealed an excellent level of consensus between the two nurses employing the AEP in all the studied departments. Overall agreement was > 92%, while the specific appropriate agreement and specific inappropriate agreement were > 88% and > 83%, respectively. External validity testing of the instrument yielded a sensitivity > 0.785, specificity > 0.55, and positive and negative predictive values > 0.775 and > 0.555, respectively. The kappa statistic for the nurses who applied the AEP and clinicians using personal judgment were perfect ($k > 0.85$) and substantial ($k > 0.68$), respectively. **Conclusions:** The results illustrate that the Iranian version of the AEP (IR-AEP) could be a reliable and valid instrument for assessing the level of inappropriate acute hospital admissions and days of care in the Iranian context.

Keywords: Appropriateness Review; Clinical Protocols; Iran; Reliability and Validity

1. Background

Internationally, the appropriate use of acute hospital beds is a main concern of policy makers and hospital practitioners (1). Health expenditures in Iran constituted 5.60% of the gross national product in 2010, and hospital expenditures increased by more than three times from 2002 to 2007. This occurred despite the fact that hospital care comprises nearly 50% of health expenditures (2). With the assumption that multiple admissions and days of care (DOC) may be inappropriate, interest in the utilization of review tools has increased (3).

Evaluating whether or not an admission or DOC is inappropriate is a difficult task, as there is no gold standard. Besides, there is no particular set of criteria that is generally applicable. One of the most widely used tool for this purpose is the Appropriateness Evaluation Protocol (AEP) developed in the United States. Its reliability and validity have been tested in various countries since its development in 1981 (4). However, because of the difference in health care delivery systems between the United States

and Iran, the original AEP had to be modified for domestic use. Modification of the original AEP by different countries has proved to be quite useful (5, 6).

The AEP is divided into a series of separate sets of admission and DOC criteria. Reviewers, employing the AEP, review the admission or DOC to decide if features in the admission or DOC fulfill any of the particular criteria (Appendix 1). Admission and the DOC are judged as appropriate if one or more of the related criteria are fulfilled.

In Iran, seven studies assessed the appropriateness of acute hospital admissions and DOC, which employed the AEP without any adaptation or modification. The prevalence of inappropriate admissions and DOC reported in the Iranian hospitals ranged from 22% to 29.6%, respectively (7). However, there are deficiencies in such studies. First, there are noticeable differences between the Iranian and the American health care systems. Therefore, the original AEP may not be applicable to Iran. Second, the reliability and validity of the AEP have not been tested in Iran.

2. Objectives

This paper aims to modify the AEP and test its reliability and validity for measuring the level of inappropriate admissions and DOC in the Iranian hospitals.

3. Patients and Methods

3.1. Cross-Cultural Translation

The original version of the AEP was translated into Persian using back-translation method (8) with three expert clinicians (an internist, a surgeon and a gynecologist), a methodologist, a language professional, and a translator. The final translation was evaluated and endorsed by the working team.

3.2. Validity and Reliability

The Persian version of the AEP was modified in a two-stage process, closely following the approach used by its American developer. A multidisciplinary expert panel of six physicians (two internists, two surgeons and two gynecologists) was selected by the research team to modify the instrument and determine its content validity using a nominal group technique. Subsequently, a retrospective descriptive study of the patients' medical records was conducted. The study was performed on a random selection of case records admitted to three different departments (internal medicine, general surgery and gynecology) of the one of the largest referral teaching hospitals (Imam Khomeini Hospital) owned by Tehran University of Medical Sciences) in Tehran, Iran. The sample size was calculated considering a disagreement degree of 20% with a two-tailed confidence interval of $\pm 5\%$ and 95% confidence. A minimum sample of 246 hospital admissions was calculated and 25% was considered to compensate the exclusion-associated losses (totally 307). The patients admitted for elective surgery, burns, intensive care, psychiatric, and patients younger than 18 years old were excluded. The data were collected from March 1 to May 31, 2013. During each day of the study, a patient's medical record was randomly selected from each department, which resulted in a total of 273 patients' records reviewed. The patients' files were summarized by one of the authors (Anvar Esmaili) using a standardized abstract format. The two reviewers were expert nurses. The nurses were trained to apply the AEP. Two nurses concurrently and independently reviewed medical records using the IR-AEP.

In the addition, assessing admission details, the two nurses also assessed the appropriateness DOC of the records with 453 DOC, where patients stayed in the hospital for longer than 24 hours (235 records).

Six clinicians (two internists, two general surgeons, and two gynecologists) were recruited to form an expert panel and also assessed the 273 admissions and 453 DOC, using their own subjective judgment about the appropriateness of the admissions and DOC. They were required

not to have direct involvement in the care of the recruited patients. Their consensus served as the gold standard with regard to the appropriateness of an acute admission and DOC. All the clinicians used the same set of records. Both groups of raters (expert clinicians and nurses) were blinded to the judgment of each other (6). To safeguard clinical file information confidentiality, the standardized abstract format was copied and patients' identifications (ID) were deleted. The physicians were identified with an anonymous ID code.

For the inter-rater reliability analysis between the pair of nurses and between the pair of clinicians, information on the hospital admissions and DOC was obtained from patients' medical records. For the intra-rater reliability, each nurse evaluated admission and DOC in each clinical file by applying the AEP instrument individually and on two occasions, separated by a 2-week interval. If each one of the criteria of the IR-AEP was fulfilled, an admission and DOC were judged as appropriate. These results were then compared to those of the expert panel.

Statistical analysis of AEP reliability was evaluated through the overall agreement. "Overall agreement is the proportion of judgments in which two reviewers agree. Specific inappropriate agreement is the proportion of judgments (among those judged to be inappropriate by at least one of the two reviewers) that are rated as being inappropriate by both reviewers" (9). Specific appropriate agreement is also calculated in a similar way. The Cohen kappa coefficient was used to assess agreement (inter-rater) above the level expected by chance (10).

The AEP validity refers to measurements in which the results obtained agree with the true result or with a gold standard. In this study, indices employed to evaluate criterion validity also included the overall agreement and the specific agreement, which included only the coincidence in agreements between the pair of nurses and the pair of clinicians. We determined AEP sensitivity and specificity, as well as a positive and negative predictive value, by considering the agreements of the clinicians' pairs as the gold standard. Also, the validity of the IR-AEP was represented by the kappa statistic.

Landis and Koch's guidelines were used to interpret levels. According to these guidelines, coefficients between 0.41 and 0.60 are regarded as moderate, between 0.61 and 0.80 as substantial, and between 0.81 and 1.00 as almost perfect (11, 12).

Statistical analysis was performed using SPSS version 16.0.

The Ethics Committee of Tehran University of Medical Sciences endorsed the study (February 23, 2013, approval No: 91/D/325/1571).

4. Results

4.1. The Consensus Process

Numbers of criteria in original AEP were transferred

to the Iranian version without any modification. Major modifications were made to the criteria in cases where a patient is admitted for electrolyte abnormalities. The specialist group asserted that additional criteria are needed for the admissions division. Accordingly, modifications regarding criteria 9 ('Electrolyte Abnormality'), 10 ('Decrease in Hematocrit'), 13 ('Refractory Hypoxemia'), 15 ('Unbearable Pain'), 16 ('Acute Abdominal Pain'), 17 ('Non-compliance with a Therapeutic Regimen'), 18 ('Discoloration of Peripheral Extremities'), and procedures that outpatient departments are not responsible for were made. For instance, a patient who may need 'Intramuscular and/or subcutaneous injections at least three times daily' is admitted to the hospital in the USA, routinely. However, such a patient may not be admitted to the hospital in Iran. Basically, most debates focused on the above mentioned criteria. Major modifications were made regarding these criteria.

As for the DOC criteria, several modifications were made regarding the physiological condition of the patients. The specialties' group demanded clarifications regarding a number of criteria in cases where there is no alternative or home care. In this regard, they recommended that a patient should not remain hospitalized for paramedical and community services, except in cases where they may need interval care.

4.2. Results of Reliability and Validity Testing

We selected 307 hospital admissions in a simple random sampling method in which 11% of the patients were excluded. Out of the 273 patient files, 732 DOC were obtained. The discharge days were excluded. Also, those files which lacked information referring to the day of the clinical file evaluation were excluded and only 453 DOC remained in the study sample (internal medicine = 206, general surgery = 132 and gynecology = 115).

During the nurses' training, an intra-reviewer agreement for hospital admission and DOC yielded kappa coefficients >0.92 and >0.94 for nurses 1 and 2, respectively. Inter-reviewer agreement on hospital admission and DOC were proved until the value of 0.88.

Although in the initial stages of this study the override was 5.7%, because overrides may be misused by inexperienced reviewers or because of bias likelihood (9), we avoided using the override option. Table 1 shows selected characteristics for all admissions.

The reliability testing results are shown in Table 2. In general, overall agreements on the hospital admissions and DOC assessments are very high (95% and 94%, respectively). Similarly, Cohen's kappa coefficients (0.88) show perfect agreement. The values obtained are highly significant ($P < 0.0001$).

Table 1. Selected Characteristics of the Study Population in Each Department ^{a, b}

	Age		Gender		Insurance		AED	
	40 \geq	40 <	Male	Female	Yes	No	Yes	No
Medical	26 (29)	65 (71)	29 (32)	62 (68)	87 (96)	4 (4)	69 (76)	22 (24)
Surgical	57 (63)	34 (37)	43 (47)	48 (53)	85 (93)	6 (7)	64 (70)	27 (30)
Gynecology	72 (79)	19 (28)	0	91 (100)	84 (92)	7 (8)	49 (54)	42 (46)

^a Abbreviations: AED, Admitted via the Emergency Department.

^b Data are presented as No. (%).

Table 2. Inter-Rater Reliability of the AEP by Departments (the Two Nurses) ^{a, b}

Reliability Measure	General Surgery (n = 91, 132) ^c	Internal Medicine (n = 91, 206) ^c	Gynecology (n = 91, 115) ^c	All Departments (n = 273, 453) ^c
Admission				
Overall agreement ^d	97	94	97	95
Cohen's K ^e	0.87	0.88	0.90	0.88 (0.87-0.90)
SAA ^d	93	92	97	94
SIA ^d	83	85	83	84
Day of Care				
Overall agreement ^d	97	92	95	94
Cohen's K ^e	0.92	0.85	0.90	0.88 (0.85-0.92)
SAA ^d	96	88	93	91
SIA ^d	90	84	88	86

^a Abbreviations: SAA, Specific Appropriate Agreement; SIA, Specific Inappropriate Agreement.

^b Average inappropriate ratings by AEP reviewers on admissions = 24.7%, and on the day of care = 34.7%.

^c n is for admission and days of care, respectively.

^d Data are presented as %.

^e Data are presented for 95% CI for K and $P < 0.0001$.

Table 3. Validity of the AEP When Compared With the Judgments of Expert Physicians ^a

	General Surgery (n = 91, 132) ^b	Internal Medicine (n=91, 206) ^b	Gynecology (n=91, 115) ^b	All Departments (n=273, 453) ^b
Admission				
Sensitivity ^c	91.5	78.5	95	92.5
Specificity ^c	97.5	55	100	90
Positive Predictive value ^c	99	77.5	100	97
Negative Predictive value ^c	80	55.5	73	79
Cohen's K ^d	0.82 (0.72-0.86)	0.74 (0.62-0.80)	0.8 (0.62-1)	0.80 (0.75-0.83)
Day of Care				
Sensitivity ^c	98	86	90	95
Specificity ^c	90	88	95	84
Positive Predictive value ^c	98.5	92	97	97
Negative Predictive value ^c	96	80	80	79
Cohen's K ^d	0.94 (0.88-0.96)	0.73 (0.67-0.78)	0.80 (0.75-0.81)	0.80 (0.76-0.84)

^a Overall agreements for admission in the General surgery department = 92.5%, Internal medicine = 88.3%, Gynecology = 95.5% with an average of 92% for all departments. Overall agreements for days of care in General surgery department = 97.25%, Internal medicine = 86.75%, and Gynecology = 91%, with an average of 91% for all departments.

^b n is for admission and days of care, respectively.

^c Data are presented as %.

^d Data are presented for 95% CI for K and P < 0.0001.

The validity of the IR-AEP was tested by comparing the assessments of nurses with personal judgments made by the expert physicians regarding appropriateness of admissions and DOC. In our study, a substantial level of agreement on admissions ($k = 0.76, 0.75 - 0.83$) and a perfect level of agreement on DOC ($k = 0.84, 0.76 - 0.86$) were obtained between members of the expert clinicians. When all raters' results on hospital admissions and DOC were combined (Table 3), the IR-AEP had a sensitivity > 0.925, specificity of > 0.84, and positive and negative predictive values > 0.97 and > 0.79, respectively. Cohen's statistic (0.80) indicated almost perfect agreement.

In general, overall agreement is always rated higher by reviewers using the objective criteria of the IR-AEP on hospital admissions and DOC (95% and 94%, respectively), as compared with 91% and 94% of clinicians using their subjective judgment. Reliability of the IR-AEP based on nurses' agreement was higher ($k=0.88$), as compared with reliability obtained from experts' subjective judgment ($k= 0.76$ and 0.86).

5. Discussion

The present study is a first attempt to customize the American AEP for the Iranian health system.

5.1. Hospital Admissions

For all departments, the level of overall agreement (95%) and specific appropriate agreement (84%) between nurses

are in close proximity with those of Sanchez-Garcia (86% and 26%, respectively) in geriatric admissions (13). The values of kappa obtained in the reliability analysis (0.87-0.90) of the tool are higher than those reported by previous investigators using the original AEP in the USA (0.44) (4). In the validity testing, the level of overall agreement between nurses and internists (88%) and between nurses and general surgeons (92.5%) is somewhat in close proximity with that reported by Sanchez-Garcia (95.5% and 94.6%, respectively). The specificity and negative predictive value of the IR-AEP achieved in this study for internists (0.55 and 0.555) and general surgeons (0.975 and 0.80) are lower and close to those reported by Sanchez-Garcia (0.96 and 0.99, respectively) (13).

5.2. Days of Care

The levels of overall agreement (94%) and specific inappropriate agreement (91%) obtained between nurses from all departments are close to and higher than those reported by the original developers of the AEP (94.3% and 79.3%, respectively) (4). Also, the values of kappa obtained in the reliability analysis of the tool (0.85-0.92) are higher than those reported by previous investigators using the original AEP in the USA (0.59-0.73) (14). The kappa coefficient for inter-rater agreement (0.88 for nurses) is higher than that reported by Kaya (0.80) (9).

In the validity testing, sensitivity (0.925), specificity (0.084), positive predictive value (0.097), and negative

predictive value (0.079) of the IR-AEP achieved in this study for hospital admissions and days of care are higher than those reported by Kaya (> 0.073 , > 0.062 , > 0.080 and > 0.073 , respectively) in all departments (9).

In this study, the values of sensitivity, specificity, positive predictive value, and negative predictive value obtained for the IR-AEP are high for evaluating hospital admissions and days of stay in all departments. However, specificity and negative predictive value in the internal medicine department for hospital admissions are moderate. Therefore, the IR-AEP has a moderate validity in terms of appropriate admissions in the internal medicine department.

When employing the IR-AEP in Iran, it is important to note that there are no replacements for acute care facilities (e.g. nursing homes, chronic care hospitals, hospices health home) in this country (15).

As it has been recommended that completion of medical records should be conducted by different health professionals, in this study, recruiting only nurses to collect data can be regarded as a limitation. Also, the retrospective design of the study may insert bias in gathering data for assessment.

The results obtained in this study show that the IR-AEP is a reliable and valid instrument for assessing appropriateness hospital admissions and DOC in Iran. It can be applied in other health care settings as well as hospitals. Considering similarities in the cultures and structures in the health services of developing countries in the Middle East, this tool could also be utilized in this region with minor modifications.

Acknowledgements

The authors would like to thank the following persons for their vital contributions to the study: Nasser Farrokh, Zine El Abidine Ebrahimzadeh, Salam Azizpour, Shahram Azizi, Masoumeh Manafpour, Goharshad Gharanzadeh, Ramin Ghaderi, Sakineh Amini. Thanks are also due to managers and personnel of Imam Khomeini Hospital affiliated to Tehran University of Medical Sciences, Tehran, Iran.

Authors' Contributions

Study concept and design: Anvar Esmaili Hamid Ravaghi and Hesam Seyedin; Analysis and interpretation of data: Hamid Ravaghi; Drafting of the manuscript: Anvar Esmaili and Bahram Delgoshaei; Critical revision of the manuscript for important intellectual content: Bahram Delgoshaei; Statistical analysis: Masoud Salehi.

Funding/Support

This study was a part of a PhD thesis supported financially by Tehran University of Medical Sciences, Tehran, Iran (grant No: 801 dated February 16, 2013).

Appendix

Adult AEP Admission Criteria-The Iranian Version

A. Severity of illness criteria

1. Sudden onset of unconsciousness or (coma or unresponsiveness) disorientation.
 2. Pulse rate:
 - Less than 50 per minute
 - Greater than 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 loss of sight or hearing.
 5. Acute loss of ability to move major body part.
 6. Acute ataxia.
 7. Persistent fever (37.8°C orally or 38.3°C rectally for > 5 days).
 8. Active bleeding, which could lead to circulatory embarrassment if homeostasis are not secured.
 9. Severe electrolyte/blood gas /CBC abnormality (any of the following, in mmol/L):
 - $156 < \text{Na} < 123$
 - $6.0 < \text{K} < 2.5$
 - $20 > \text{CO}_2$ combining power > 36 (unless chronically abnormal)
 - Arterial $7.45 < \text{pH} < 7.30$
 - BUN > 45 mg/dL (with increased Cr)
 - FBS ≤ 70
 - FBS ≥ 400 (associated with symptoms of weight reduction for no apparent reason, polyuria, and polydipsia)
 - WBC ≥ 12000 (ESR ≥ 30)
 - WBC ≥ 2500 (ANC ≥ 1000).
 10. Hematocrit $< 30\%$ (in cases where it cannot be handled by outpatient services or has originated from an underlying disease).
 11. Acute or progressive sensory, motor, circulatory or respiratory embarrassment that can incapacitate the patient (inability to move, feed, breathe, urinate, etc.).
 12. ECG evidence of acute ischemia (must be suspicious of a new MI).
 13. Refractory hypoxemia with $\text{SpO}_2 < 85\%$ and $\text{RR} \geq 40$ (unless chronically abnormal).
 14. Wound dehiscence or evisceration.
 15. Unbearable pain (that may not be managed by outpatient services (including A and E and GP) management.
 16. Acute abdominal pain with laboratory, ultrasonography, X-ray abnormality that has lasted over 6 hours.
 17. Noncompliance with a therapeutic regimen where failure to comply could seriously endanger the patient's health (where/when there is no alternative or a layman trained to do this at home).
 18. Severe pain, sudden swelling, discoloration with cold peripheral extremities and decrease in distal pulse.
- ### B. Clinical Service
1. Parenteral therapy-Intermittent or continuous IV fluid replacement and/or nutrition (does not include tube feedings).
 2. Surgery or procedure scheduled on the same day, requiring:

-General or local anesthesia or

-Equipment or facilities available only for inpatients.

3. Vital sign monitoring every 6 hours or longer (may include telemetry or cardiac monitor), which is beyond the scope of management of an emergency ward.

4. Chemotherapeutic agents that require continuous observation for life-threatening toxic reaction.

5. Intravenous medication at least every 8 hours provided that the physician switches to oral medications within 24–48 hours.

6. Intermittent or continuous use of a respirator at least every 6 hours.

Adult AEP Day of Care Criteria, the Iranian Version

A. Medical Services

1. Procedure in the operating room on the same day, requiring clinical control or observation.

2. Scheduled procedure in the operating room for the next day, requiring preoperative consultation or evaluation, followed by preoperative preparation requiring hospital facilities/personnel 24 hours (48 hours for bowel surgery) prior to the operation.

3. Cardiac catheterization on the same day.

4. Angiography on the same day.

5. Biopsy of internal organ on the same day.

6. Thoracentesis or paracentesis on the same day.

7. Invasive CNS diagnostic procedure (e.g., lumbar puncture, external tap, ventricular taps, and pneumoencephalography) on the same day.

8. Outside-hospital investigation on the same day.

9. Any test requiring strict dietary control, for the duration of the diet (where/when a layman has not been trained to do this at home).

10. Referral to other specialty unit on the same day.

11. Dose adjustment in potentially serious condition.

12. Essential close medical monitoring by a medical officer at least three times daily (observation must be documented in the record).

13. Postoperative day for any procedure explained above (1 and 3-7), 4 days for abdominal surgery, head, neck, limb, and mastectomies, 1 day for appendectomies, hernias, hydroceles, and biopsies, and 0 days for cardiac catheterization, angiography, thoracentesis, and paracentesis.

B. Nursing/Life Support Services

1. Respiratory care -intermittent or continuous use of a respirator and/or inhalation therapy (with chest PT, IPPB) at least thrice daily.

2. Parenteral therapy -intermittent or continuous IV fluid or nutrition (where/when a layman has not been trained to do this at home).

3. Continuous vital sign monitoring by nurses -at least every 30 minutes, for at least 2 hours, specifically requested by the doctor.

4. Intravenous medication at least thrice daily provided that the physician switches to oral medications within 24–48 hours (where/when a layman has not been trained to do this at home).

5. Intake and output measurement required by the doc-

tor (where/when a layman has not been trained to do this at home).

6. Surgical wound and/or drainage care (chest tubes, T-tubes, hemovacs, Penrose drains).

7. Close medical monitoring by a nurse, at least three times daily, under a doctor's orders.

8. Interval cares where there is no other alternative.

C. Patient Condition

-Within the past 24 hours:

1. Inability to void or move the bowels (past 24 hours) not attributed to neurologic disorder.

2. Potentially serious condition suspected-for observation (e.g. Head injury).

3. Transfusion due to acute blood loss.

-Within the past 48 hours:

1. Ventricular fibrillation or ECG evidence of acute ischemia, as stated in progress note or in ECG report.

2. Fever > 37.8°C orally or 38.3°C rectally, if the patient was admitted for reasons other than fever.

3. Coma-unresponsiveness for at least one hour.

4. Acute confusional state, not due to alcohol intoxication.

5. Acute hematologic disorders, significant neutropenia, anemia, thrombocytopenia, leukocytosis, erythrocytosis, and thrombocytosis yielding signs or symptoms.

6. Progressive acute neurologic deficits.

-Within 5 days before day of review:

1. Documented occurrences-new acute myocardial infarction or cerebrovascular accident (stroke).

References

- Kossovsky MP, Chopard P, Bolla F, Sarasin FP, Louis-Simonet M, Allaz AF, et al. Evaluation of quality improvement interventions to reduce inappropriate hospital use. *Int J Qual Health Care.* 2002;**14**(3):227-32.
- Iran SCI. Iran National Health Accounts; Annual Report, 2002 to 2007. In: Statistical Center of Iran, editor. Tehran: 2011.
- Esmail A. Development of the Paediatric Appropriateness Evaluation Protocol for use in the United Kingdom. *J Public Health Med.* 2000;**22**(2):224-30.
- Gertman PM, Restuccia JD. The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. *Med Care.* 1981;**19**(8):855-71.
- Panis LJ, Verheggen FW, Pop P. To stay or not to stay. The assessment of appropriate hospital stay: a Dutch report. *Int J Qual Health Care.* 2002;**14**(1):55-67.
- Leung LP, Fan KL. Who should be admitted to hospital? Evaluation of a screening tool. *Hong Kong Med J.* 2008;**14**(4):273-7.
- Hatam N, Askarian M, Sarikhani Y, Ghaem H. Necessity of admissions in selected teaching university affiliated and private hospitals during 2007 in Shiraz, Iran. *Arch Iran Med.* 2010;**13**(3):230-4.
- Guile R, Leux C, Paille C, Lombrail P, Moret L. Validation of a tool assessing appropriateness of hospital days in rehabilitation centres. *Int J Qual Health Care.* 2009;**21**(3):198-205.
- Kaya S, Vural G, Eroglu K, Sain G, Mersin H, Karabeyoglu M, et al. Liability and validity of the Appropriateness Evaluation Protocol in Turkey. *Int J Qual Health Care.* 2000;**12**(4):325-9.
- Paille-Ricolleau C, Leux C, Guile R, Abbey H, Lombrail P, Moret L. Causes of inappropriate hospital days: development and validation of a French assessment tool for rehabilitation centres. *Int J Qual Health Care.* 2012;**24**(2):121-8.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;**33**(1):159-74.
- Hammond CL, Phillips MF, Pinnington LL, Pearson BJ, Fakis A. Ap-

- appropriateness of acute admissions and last in-patient day for patients with long term neurological conditions. *BMC Health Serv Res.* 2009;**9**:40.
13. Sanchez-Garcia S, Juarez-Cedillo T, Mould-Quevedo JF, Garcia-Gonzalez JJ, Contreras-Hernandez I, Espinel-Bermudez MC, et al. The hospital appropriateness evaluation protocol in elderly patients: a technique to evaluate admission and hospital stay. *Scand J Caring Sci.* 2008;**22**(2):306-13.
 14. Siu AL, Sonnenberg FA, Manning WG, Goldberg GA, Bloomfield ES, Newhouse JP, et al. Inappropriate use of hospitals in a randomized trial of health insurance plans. *N Engl J Med.* 1986;**315**(20):1259-66.
 15. 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**(3):239-44.