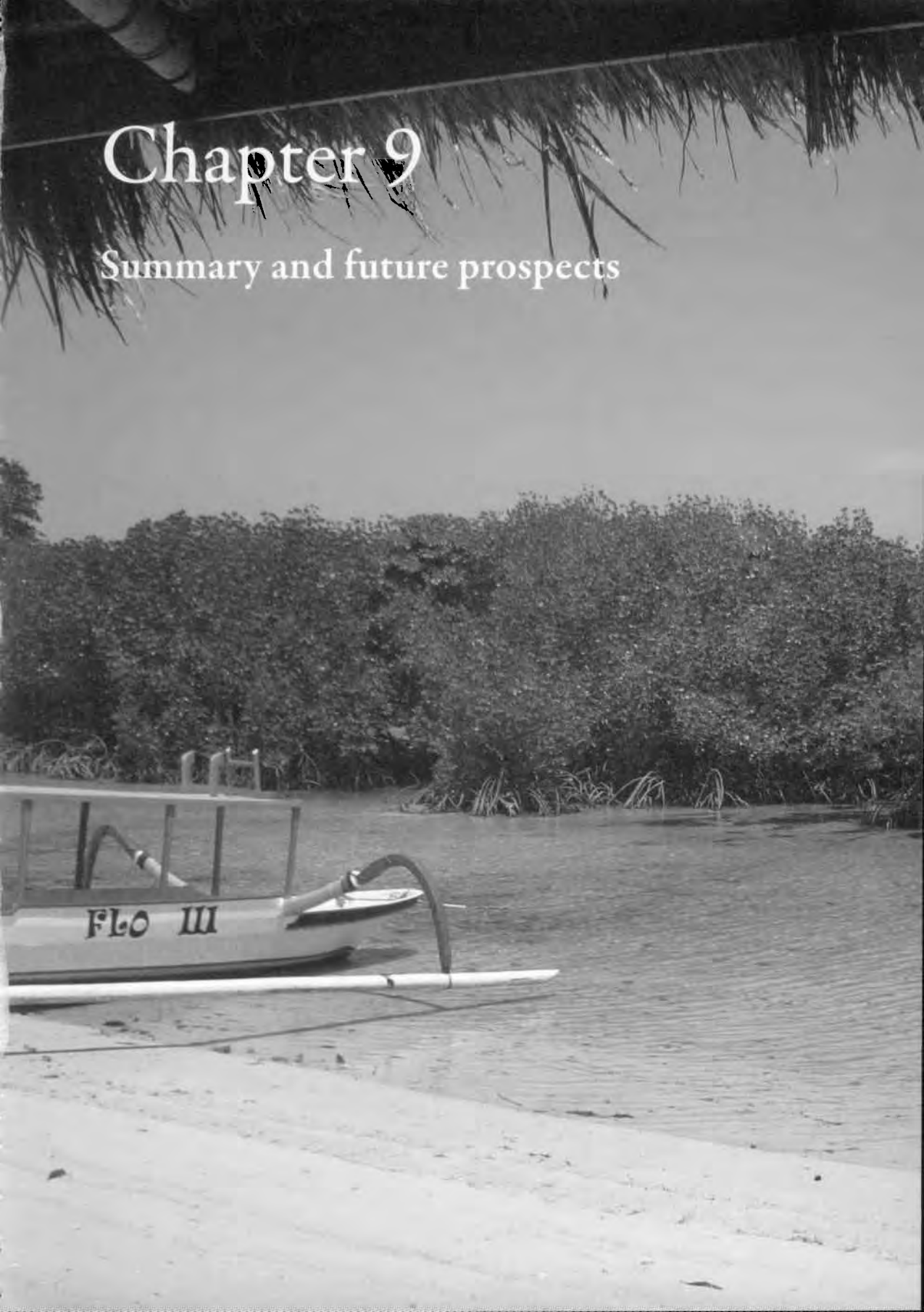
Chapter 9

Summary and future prospects



SUMMARY

In **Chapter 1** we provided an overview of the literature on the reliability and validity of triage systems in paediatric emergency care.
The Manchester Triage System (MTS), the Emergency Severity Index (ESI), the Paediatric Canadian Triage and Acuity Score (paedCTAS) and the Australasian Triage Scale (ATS) are common used triage systems and contain specific parts for children.
We concluded that the MTS and paedCTAS both seem valid to triage children in paediatric emergency care. The internal validity is moderate for the MTS and confirmed for the CTAS, but not studied for the most recent version of the ESI, which contains specific fever criteria for children. Reliability of the MTS is good, moderate to good for the ESI and moderate for the paedCTAS. More studies are necessary to evaluate if one triage system is superior over other systems when applied in emergency care.

We evaluated reliability and validity of the MTS in paediatric emergency care. The studies were performed at the Erasmus MC-Sophia Children's hospital and the Haga hospital Juliana Children's hospital. In **chapter 2** the reliability of the MTS was studied. We performed an inter-rater agreement study in which all ED nurses triaged twenty written case scenarios and actual patients were triaged simultaneously by two nurses using the MTS. The inter-rater agreement was good to very good (weighted kappa of 0.83, 95% C.I.0.74–0.91) using case scenarios and good (weighted kappa of 0.65, 95% C.I. 0.56–0.72) when actual patients were triaged.
In chapter 3 the validity of the MTS was evaluated in a large prospective observational study in 17,600 children in 2006/2007. The MTS was applied to patients attending the ED. Data was gathered to determine the urgency level according to an independent predefined reference standard for urgency. This reference standard was based on abnormal vital signs at presentation to define urgency 1, potentially life threatening condition for urgency 2, and a combination of diagnostic resources, therapeutic interventions, hospitalization and follow-up for urgency 3, 4 and 5.
Overall compliance to the MTS was 95%. The Manchester urgency level agreed with the reference standard in 4,582 of 13,554 (34%) children. In 7,311 (54%) the MTS attributed a too high urgency level (over-triage) and in 1,661 (12%) a too low urgency level.(under-triage) The sensitivity of the MTS to correctly identify high urgent patients was 63% (59 to 66) and specificity to correctly identify low urgent patients was 79% (95% CI 79 to 80).

The likelihood ratio was 3.0 (95% CI 2.8 to 3.2) for high urgency and 0.5 (0.4 to 0.5) for low urgency; though the likelihood ratios were lower for those presenting with a medical problem (2.3, 95% CI 2.2 to 2.5) versus 12.0, 95% CI 7.8 to 18.0, for trauma and in younger children (2.4, 95% CI 1.9 to 2.9) at 0–3 months v 5.4 (95% CI 4.5 to 6.5) at 8–16 years.
We concluded that the MTS has a moderate validity in paediatric emergency care. It errs on the safe side, with much more over-triage than under-triage compared with an independent reference standard. Triage of patients with a medical problem or of younger children is particularly difficult.
In **chapter 4** we evaluated patients who were severely under-triaged, compared to the reference standard for urgency. We performed a case study to determine the severity of under-triage in children. Under-triage was defined as patients triaged as low urgent (level 3-5) by the MTS and high urgent (level 1 or 2) by the reference standard, with at least two levels difference between the MTS and the reference standard urgency. Three experts in paediatric emergency care discussed cases, to determine severity. Secondly, to assess predictors of under-triage a univariate and multivariate regression analyses were performed. Under-triage could be considered as severe in 70% (N=107/152, 8 cases with missing data) of the under-triage. The reference standard was in 78% (N=83) of those patients determined by abnormal vital signs. Further, children younger than four years of age and children assigned to the MTS flowchart 'unwell child' are more likely to be under-triaged than children assigned to other flowcharts. ($OR_{<3\ months}$ 7.2, 95% CI 4.3–12.1, $OR_{3\text{-}11\ months}$ 2.6, 95% CI 1.5–4.6, $OR_{<1\text{-}4\ years}$ 1.8, 95% CI 1.1–3.0 and $OR_{unwell\ child}$ 5.6, 95% CI 2.6–11.9). Under-triage might have serious consequences in a few patients. The validity of the MTS may improve by adding abnormal vital signs as a discriminator in young children and in the MTS flowchart 'Unwell child'.
In **chapter 5** we studied the value of body temperature combined with age and presenting problem to predict high urgency, according to the reference standard for urgency.
Temperature separately had a moderate discriminative ability to predict urgency (AUC 0.58, 95% CI 0.56–0.60), but in combination with presenting problem and age the performance improved.(AUC 0.75, 95% CI 0.73–0.76) Temperature influences urgency especially in patients presenting with upper respiratory tract and urinary tract problems. We concluded that body temperature, combined with age and presenting problem is an important discriminator in triage systems. Together, these discriminators contribute to differentiate the triage decisions and can be implemented in different triage systems.

We modified the MTS for patients with fever based on age and presenting problem and for other specific patient groups, such as for patients with only a recent problem as discriminator,

and implemented the modified MTS at both EDs. In **chapter 6** we evaluated the external validity of the modified MTS in 11,481 patients by comparing the modified MTS urgency level to the reference standard urgency in both hospitals in 2007/2008. Compared to the original MTS specificity improved from 79%, 95% C.I. 79 to 80% to 87%, 95% CI 86 to 87% while sensitivity remained similar (63%, 95% CI 59 to 66%) versus (64%, 95% CI 60 to 68%). The diagnostic odds ratio increased, from 4.1 (95% CI 3.2 to 5.1) to 11 (95% CI 9.6 to 14).
We concluded that Modifications of the MTS for paediatric emergency care resulted in an improved specificity while sensitivity remained unchanged. Further research should focus on the improvement of sensitivity.

In the final section (**chapter 7 and 8**) we focussed on the ability of the MTS to identify low urgent patients in order to refer these patients to another healthcare professional.
In **chapter 7** we assessed hospitalization, as a proxy for safety and determinants for hospitalization for low urgent, self-referred patients (MTS level 4 or 5) presenting at the ED. Secondly, discharged patients received a telephonic follow-up 2-4 days after consultation.
Among 5,425 patients, 191 (3.5%) were hospitalized. Hospitalization was more likely for children younger than one year of age (OR 3.0, 95% CI 2.2 to 4.1) and for patients presenting with dyspnea (OR 2.5, 95% CI 1.5 to 4.1) gastrointestinal problems (OR 3.5, 95% CI 2.5 to 4.9) and for patients with fever without other specific symptoms. (OR 2.8, 95% CI 1.1 to 7.2). 3,975 / 5,234 (76%) could be contacted for follow-up after discharge. After ED discharge only six (0.15%) patients were hospitalized.
Referral of low urgent, self referred children to another healthcare professional may be safe except for children aged under one year or when presenting with dyspnea, gastrointestinal problems and for patients with fever without specific symptoms.

In **chapter 8** we evaluated compliance and effects on costs when low urgent, self referred children, who visited the ED were actually referred to the general practitioner cooperative.
During six months 140 patients were referred to the general practitioner cooperative. 101/140 patients (72%) were reached during telephonic follow up. After discharge seven patients (7%) had an unscheduled revisit. No patients were subsequently hospitalized. Patient satisfaction was graded as 6.6 (95% CI 6.2–7.1).
275 patients were included to study compliance. 95/247 (38%) patients were referred to the GP. 46/247 parents (19%) refused referral. For 106/247 patients (43%) referral was not initiated by the nurse due to co-morbidity or the nurse expected she could not convince the

parents. Data on 28/275 patients (10%) were missing. Mean costs per low urgent patient were €106, when initially seen at the ED and €101 after implementation of GP referral. Larger cost reductions are feasible if more patients are referred and patients would be referred during daytime as well.

We concluded that parents and children were moderately satisfied and referral resulted in a small cost reduction. Effectiveness was not optimal since a minority could be referred and many patients refused referral.

FUTURE PROSPECTS

Validity of triage systems

In order to evaluate the validity of a triage system a reference standard for urgency should be defined. In the past different methods were used. Trends in resource use and hospitalization in relation to the urgency classification were studied in several observational studies. In smaller studies an expert panel defined the reference urgency classification.[1]
The aim of triage is to determine the urgency of the patient. Urgency is based on presenting symptoms and partly determined by the patient's working diagnosis. However, urgency might differ between patients with the same diagnosis. For example not all patients with a serious bacterial infection will need very urgent care. Patients with pneumonia with normal vital signs will need a lower urgency level than a patient who present with a septic shock.

Triage should determine urgency at the time of triage. Especially for urgent presentations, the patient's condition may change quickly over time. That's why it is important to use a reference standard based on the patient's condition which has been measured within a short time frame from the triage moment.[2]

The reference standard we used determines urgency based on a number of items. Deviated vital signs at presentation defined an 'immediate' level. A potentially life threatening condition defined a 'very urgent' level, as stated at the end of ED consultation. The three lowest urgency levels ('urgent', 'standard' or 'non urgent') were defined based on the amount of resources used (diagnostics, treatment, hospitalization) and scheduled follow up.
Although all items of the reference standard are related to urgency we are aware that they do not precisely define urgency. By combining the items we developed a more precise measure to determine five urgency levels.

Improvements of the Manchester Triage System for children

With specific modifications, mainly for children with fever, based on the presenting problem and age we further improved specificity to identify true non-urgent patients based on the reference standard. The modifications did not improve sensitivity of the system to identify true high urgent patients, it remained 63%. Sensitivity focuses on the two highest MTS urgency levels. Our study on the validity of the original MTS showed that these two levels account for only 5.2% of the population following the reference standard for urgency. (Chapter 3)

When studying high urgent patients who were not correctly triaged (severe under-triage), we showed that especially patients with deviated vital signs were severely under-triaged. The MTS identifies 'Immediate' patients by describing conditions in which care should be delivered immediately. These conditions described the consequences of severe deviation of vital signs such as airway compromise, inadequate breathing and shock. Patients with only elevated heart rate, deviated blood pressure or irregular heart rhythm, will be triaged into a lower urgency levels based on other present discriminators.
Further modifications should be studied and focus on the inclusion of vital signs into the MTS in order to identify high urgent patients.[3]

By comparing the MTS urgency levels with the reference standard urgency we could identify discriminators, which showed to have a better validity when linked to a higher or lower urgency. Using this strategy we studied patient groups triaged with the ten most common used MTS flowcharts. These flowcharts accounted for 80% of the patients with non-traumatic problems. We were limited to study only the common used discriminators within these flowcharts such as 'Recent Problem' (20%), pain discriminators (17%), fever discriminators (15%), 'Recent Injury' (9%), 'Increased work of breathing' (4%) and 'Persistent vomiting' (4%) (Chapter 3).
With more extended data collection we may study less common used discriminators in order to identify more patient groups in which validity of the MTS is low, aiming to further improve the MTS.

In the presented study we only studied and modified discriminators which are present in the original MTS. Based on comments from users and literature specific discriminators could be added and studied. New discriminators such as seasonality, comorbidity and more specific discriminators will be likely to further improve the MTS. However, a large dataset is necessary to have sufficient power. A multicenter study could result in a larger number of included patients. Compared to EDs in the UK and US, the patient load visiting the EDs in the Netherlands is relatively low. Including EDs from the UK could increase generalizability and efficiency.

In the second version of the MTS as proposed by the Manchester Triage Group, some minor modifications, which were based on comments from users, were inserted.[4] The flowchart 'Unwell Child' was changed and contains new discriminators as 'Fails to react to parents'

and 'Signs of Meningism' which lead to 'Very Urgent' category. Before, children were triaged using the 'General flowchart', which did not contain specific discriminators for children.[5] We applied and studied the original MTS. The second version combined with the described modifications, is now used at the ED of the Haga hospital, Juliana Children's hospital and the Erasmus MC-Sophia Children's hospital. It is important to externally validate the modified second version of the MTS in a new population.

Methodology

Triage systems are based on consensus based decision rules. To validate triage systems the methodology of diagnostic research can be applied.

However, some specific factors of triage classification differ from clinical decision models, and should be considered. We used multivariate logistic regression modelling to study temperature as discriminator in triage systems. (Chapter 5) We combined the categories of the five level reference standard for urgency, into two categories. The categorization of the reference standard leads to a more simplified final result. We studied the risk of high urgency, and therefore did not further differentiate between the two highest urgency categories and the three lowest urgency categories. Further research could apply multinomial, or proportional odds regression analysis, for which ordinal variables, as the five level reference standard for urgency can be used as outcome measure.

When studying options for modifications the risk of increasing over- or under-triage should be taken into account. Sensitivity and specificity express the balance between over- and under-triage. A five level triage system is categorized into the two highest and three lowest urgency categories. Experts can decide if improving sensitivity is more important than improving specificity or the other way around. The value of over-, under- and correct triage is based on the number of categories over- or under-triage compared with the reference standard. By comparing the triage urgency level to a reference standard urgency level, weights can be assigned for the number of categories over- or under-triage for different urgency levels. In the literature some suggestions for weighting were proposed.[6,7] Under-triage is weighted as more severe than over-triage. They can be used to further study validity of triage systems.

The aim of triage is to see patients first who will be harmed if the initiation of treatment is delayed. To reach this aim, specific discriminators are needed which can correctly identify patients with high urgent conditions. The patient group who presents at the ED represents a wide range of different problems. Secondly, the triage assessment should be very short in order

not to delay treatment by the triage process itself, so only easily identifiable discriminators can be used for a triage assessment. It follows that highly specific discriminators which can be used for triage do probably not exist. A triage system with a high sensitivity and high specificity is therefore probably impossible to develop. A system with a high sensitivity will optimally identify high urgent cases. However, the consequence of a low specificity will result in many patients who have to be seen within a very short time frame. If the ED capacity is not sufficient, all high urgent patients (true-positive and false-positive), have to wait longer than their maximum time frame. A high specificity with a low sensitivity will result in more false-negative cases.

An optimal balance between the number of false-negative and false-positive classifications can be determined in discussion with experts. In chapter 4 we showed that the under-triaged patients (the false-negative patients) may result in severe consequences in 70% of the cases, according to the experts. Actual effects of under- and over-triage are hard to study since many factors more than the triage process determine morbidity and mortality.

Furthermore, a formal decision-analytic perspective can be used. A relatively simple approach is to consider the 'net benefit' of a triage system.[8,9] The net benefit is a weighted sum of true-positive and false-positive classifications, where the relative weight of false-positive classifications is given by the odds of the decision threshold to define an urgent versus a non-urgent case. With a low threshold, the relative weight is low, and true-positive classifications are far more important than false-positive classifications. The net benefit calculation indicates whether the model is beneficial in terms of clinical consequences, compared to treating all patients as high urgent or all patients as low urgent.

A more extensive approach is a formal cost effectiveness study. Costs of the effect of a longer waiting time on short and long term consequences as discussed by the experts can be calculated.

When a triage system is not sensitive and specific enough, more physicians should be hired in order to see patients within a sufficient time frame. However, we have to take into account that presentation of level 1 urgency patients, who require very time consuming care, is difficult to predict based on historical data, since they do not present often. A simulation model could possibly be developed to determine optimal time frames in which patients are

seen, based on distribution of urgency levels, time of presentation, the aimed sensitivity and specificity of the triage system along with the additional costs.

Low urgent patients at the emergency department

In chapter 7 we studied safety of the MTS to identify low urgent patients. We estimated that referral for specific patient groups will be safe; depending on the proportion of patients who are hospitalized when they consult the ED. Safely identified low urgent patients could be seen by another caregiver such as a general practitioner.

In chapter 8 compliance and effects on costs when MTS low urgent patients are actually referred to the general practitioner cooperative (GPC) were evaluated. Referring low urgent self referred children to a GPC resulted in a small cost reduction, while patient were satisfied, but compliance of referral was low.

The sample size of this study was too small to detect effects of referral on hospitalizations, as a proxy for safety. Larger studies should be performed comparing proportion hospitalization when low urgent patients consult the ED and when they are referred to the GPC.

Several GPCs are now located next to the ED and some have one entry for all patients. In this way a triage system can advice patients if they should go to the ED or GPC.

However, for this purpose it is unclear which triage system is valid and effective. We studied the MTS for its ability to identify low urgent patients for the ED setting. Patients presenting to the GPC will have a lower prevalence of conditions that require urgent consultation. For this patient group, triage criteria should be less conservative.

A new triage system, the Netherlands Triage System (NTS) was developed and based on the MTS, the Dutch National Telephone guidelines and a Dutch protocol aiming to guide pre-hospital transportation. The aim of the system was to correctly triage patients at the ED and GP setting and to provide an advice on which caregiver the patient should consult. The system was studied in a small data set during the implementation process of the system.[10] It was shown to be reliable but many patients triaged as low urgent were hospitalized (ED setting) or referred to the ED (GP setting). The power of the study was not sufficient to confirm neither reliability nor validity of the system when applied to children.

Further research should focus on the validity and improvements of this system as well, and should compare the validity of the MTS with the NTS, to see which system is superior to use at the ED setting and for the combined GPC/ED setting.

REFERENCES

1. van der Wulp I, van Baar ME, Schrijvers AJ. Reliability and validity of the Manchester Triage System in a general emergency department patient population in the Netherlands: results of a simulation study. Emerg Med J 2008;25(7):431-4.
2. Maconochie I, Dawood M. Manchester triage system in paediatric emergency care. Bmj 2008;337:a1507.
3. Thompson MJ, Coad N, Harnden A, Mayon-White R, Perera R, Mant D. How well do vital signs identify children with serious infections in paediatric emergency care? Arch Dis Child 2009.
4. Mackway-Jones K, Marsden J, Windle J. Emergency Triage, Manchester Triage Group. Second edition ed: Oxford: Blackwell Publishing Ltd; 2006.
5. Mackway-Jones K. Emergency Triage, Manchester Triage Group: London: BMJ Publishing Group; 1997.
6. Lee A, Hazlett CB, Chow S, Lau F-l, Kam C-w, Wong P, et al. How to minimize inappropriate utilization of Accident and Emergency Departments: improve the validity of classifying the general practice cases amongst the A&E attendees. Health Policy 2003;66(2):159-168.
7. van der Wulp I, van Stel HF. Adjusting weighted kappa for severity of mistriage decreases reported reliability of emergency department triage systems: a comparative study. J Clin Epidemiol 2009.
8. Steyerberg EW, Vickers AJ. Decision curve analysis: a discussion. Med Decis Making 2008;28(1):146-9.
9. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565-74.
10. van Veen M, Huibers AMJ, Giesen PHJ, Moll HA. Reliability and validity of the Netherlands Triage System at the emergency department and the general practitioner cooperative. Submitted 2009.