Predicting Hospital Admission for Emergency Department Patients using a Bayesian Network

Jeffrey Leegon, BS[1], Ian Jones, MD[2], Kevin Lanaghan, BS[3], Dominik Aronsky, MD, PhD[2,3]
Dept. of Computer Science, University of Birmingham, Birmingham, UK[1];
Dept. of Emergency Medicine[2], Dept. of Biomedical Informatics[3],
Vanderbilt University, Nashville, TN,

Hospital admission delays in the Emergency Department (ED) reduce volume capacity and contribute to the nation's ED diversion problem. This study evaluated the accuracy of a Bayesian network for the early prediction of hospital admission status using data from 16,900 ED encounters. The final model included nine variables that are commonly available in many ED settings. The area under the receiver operating characteristic curve was 0.894 (95% CI: 0.887-0.902) for the validation set. The system had high accuracy and may be used to alert clinicians to initiate admission processes earlier during a patient's ED encounter.

Background: For more than a decade, overcrowding in Emergency Departments (ED) has constituted a national crisis. EDs are the nation's safety net for access to care and have a legal mandate to see any patient seeking care. One of the primary drivers of ED diversion is a shortage of available hospital beds. This results in extended ED stays for patients who are waiting in the ED to be admitted to the hospital ("boarding"). An increasing number of boarding patients leads to limited ED capacity and restricted access for patients seeking emergency care. Because earlier initiation of admission processes may lead to a decreased burden of ED diversion, we developed and evaluated a hospital admission prediction instrument. The instrument uses only data elements that are usually available in real-time in many ED settings.

Methods: The adult ED at Vanderbilt University Medical Center is an academic, urban Trauma Level I center with an annual patient volume of >42,000 patients. The ED staff uses a computerized whiteboard for patient tracking and workflow management. To develop a probabilistic system (Bayesian network) for predicting hospital admission status, we used ED whiteboard data that were available for all patients early during their ED encounter. For the development phase of the Bayesian network, we included data from all 9,520 consecutive adult patients who had an ED encounter between 4/10/04 and 6/30/04. We excluded patients who died in the ED, left the ED without being seen or left against medical advice. We randomly assigned cases to a training (70%) and a test set (30%). To identify network structure, we developed different models using machine learning approaches and expert knowledge. The final model was validated against an independent data set that included data from 7,447 consecutive ED patients (7/1/2004-8/31/2004). We allowed information to be missing in both the training and validations sets. The area under the receiver operating characteristic curve (AUC) was used to evaluate the system's accuracy. We emphasized sensitivity over specificity; therefore, we evaluated the model's performance characteristics (specificity and predictive values) at fixed sensitivity levels of 90% and 95%.

Results: The hospital admission rate was 29% in both the development and validation set. The final network model included age (in deciles), time of registration (4-hour periods), mode of arrival (ambulance, helicopter, car, other), ICD-9 coded chief complaint, the acuity level (Emergency Severity Index), type of consult, the presence of at least one lab test, at least one radiology test and an ECG exam. Complete data were available for six out of nine variables; three variables included missing information that ranged from 1.3% (age) to 4.3% (acuity). The AUC for the training set was 0.920 (95% CI: 0.913-0.927), for the test set 0.894 (95% CI: 0.882-0.906), and for the validation set 0.894 (95% CI: 0.887-0.902). Table 1 lists the model's test characteristics at fixed 90% and 95% sensitivity levels for the independent validation set.

Table 1: Test characteristics at fixed sensitivity levels

Sensitivity (fixed)	Specificity	Predictive value pos.	Predictive value neg.
90%	71%	56%	95%
95%	62%	50%	97%

Discussion: The Bayesian network demonstrated high accuracy. The network includes variables that support its integration with an existing information technology environment for real-time admission prediction and may help initiating admission processes earlier during a patient's ED encounter.

Acknowledgment: The first author was supported by NLM Training Grant LM07450-02.