European Journal of Internal Medicine xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

European Journal of Internal Medicine

journal homepage: www.elsevier.com/locate/ejim

Original Article

External validation of a decision tree early warning score using only laboratory data: A retrospective review of prospectively collected data

Tara E. Holm Atkins[a,*], Malin C. Öhman[a], Mikkel Brabrand[a,b]

[a] Department of Emergency Medicine, Hospital of South West Jutland, Esbjerg, Denmark
[b] Department of Emergency Medicine, Odense University Hospital, Denmark

ARTICLE INFO

Keywords:
Early Warning Score
Hospital admission
In-hospital mortality
External validation study
Decision tree
Acute medicine
Clinical/laboratory parameters
Deteriorating patients
Illness severity
Internal medicine
Secondary teaching hospital

ABSTRACT

Introduction: Early warning scores (EWS) have been developed to identify the degree of illness severity among acutely ill patients. One system, The Laboratory Decision Tree Early Warning Score (LDT-EWS) is wholly laboratory data based. Laboratory data was used in the development of a rare computerized method, developing a decision tree analysis. This article externally validates LDT-EWS, which is obligatory for an EWS before clinical use.

Method: We conducted a retrospective review of prospectively collected data based on a time limited sample of all patients admitted through the medical admission unit (MAU) on a Danish secondary hospital. All consecutive adult patients admitted from 2 October 2008 until 19 February 2009, and from 23 February 2010 until 26 May 2010, were included. Validation was made by calculating the discriminatory power as area under the receiver-operating curve (AUROC) and calibration (precision) as Hosmer-Lemeshow Goodness of fit test.

Results: A total of 5858 patients were admitted and 4902 included (83.7%). In-hospital mortality in our final dataset (n = 4902) was 3.5%. Discriminatory power (95% CI), identifying in-hospital death was 0.809 (0.777–0.842). Calibration was good with a goodness-of-fit test of X^2 = 5.37 (7 degrees of freedom), p = 0.62.

Conclusion: LDT-EWS has acceptable ability to identify patients at high risk of dying during hospitalization with good precision. Further studies performing impact analysis are required before this score should be implemented in clinical practice.

1. Introduction

Early warning scoring systems were developed to ensure the timely identification, at the bedside, of patients exhibiting physiological signs compatible with established or impending critical illness [1]. Most systems only use vital signs but some use a combination of vital signs and other physiological parameters (e.g. blood tests) [2]. To present day, only a few systems using only blood tests (3, 4) have been introduced.

One system, using only blood tests, is the Laboratory Decision Tree Early Warning Score (LDT-EWS) [4]. It was developed in the UK in 2013 and utilizes data from the seven most commonly used blood tests: hemoglobin (Hgb), white cell count (WCC), urea (U), albumin (Alb), creatinine (Cr), sodium (Na) and potassium (K). A complex decision tree analysis led to the development of the score and initial validation at development found the score predicted mortality satisfactually when used as an aggregate score and not calculating absolute mortality risk.

There is uncertainty of which scores are most reliable and useful, and this can only be evaluated by impact analyses [5,6]. And while many EWSs exist, only a few have been validated and of these, only a few have been performed in depth. External validation of any risk stratifical tool to be used in clinical practice is at least highly desirable, if not obligatory, to prove validity in a relevant setting [7]. The aim of this article is to validate LDT-EWS, to see if it is applicable to other populations than the development cohort and therefore add another stone before it can be implemented in clinical practice.

2. Methods

2.1. Setting

The validation cohort is from the medical admission unit (MAU) at Hospital of South West Jutland, a 460-bed regional teaching hospital in western Denmark, with a mixed urban and rural contingency population of 220,000. The MAU has 24 beds and 10,950 adult admissions per year. It serves all sub-specialties of internal medicine. Patients can be admitted by their primary care physician, out-of-hours emergency medical service, outpatient clinics, emergency department and

* Corresponding author.
E-mail address: tara_atkins@hotmail.com (T.E. Holm Atkins).

https://doi.org/10.1016/j.ejim.2017.12.008
Received 4 February 2017; Received in revised form 6 December 2017; Accepted 18 December 2017

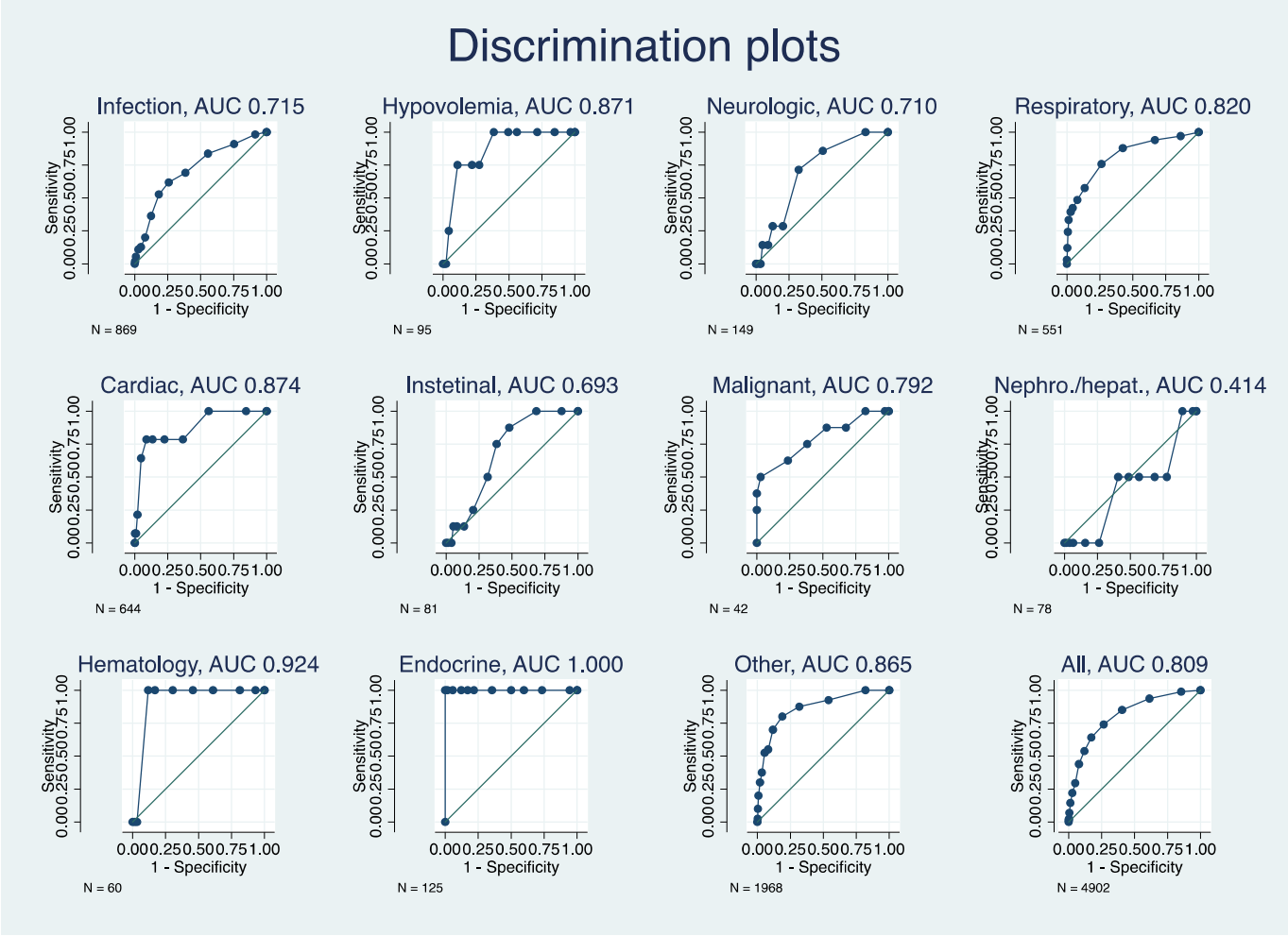

Fig. 1. Area Under the Receiver Operating Curve (AUROC) for all included patients, and for each Major Disease Categories (MDC).

ambulance service.

2.2. Design and data

We conducted a retrospective review of prospectively collected data based on a time limited sample of all patients admitted through the MAU. All consecutive adult patients (ages ≥ 15 years) admitted from 2 October 2008 until 19 February 2009, and from 23 February 2010 until 26 May 2010, were included. The data set is two consecutive periods, because it was required for another study at the time [1].

Most patients had a standard blood panel taken at admission, depending on their symptoms. No extra biochemical analyses were added as part of this study, and only blood tests ordered by the admitting doctor were included. Those who did not acquire blood tests were not included in the analysis due to missing data.

The outcome was in-hospital mortality, as in the original article. The results of blood tests were extracted from the hospital computer systems after the inclusion was completed and all patients were discharged or dead [3]. For a more thorough description of the study setup, see Brabrand et al. [3]. Data on outcome was extracted from the Danish Civil Person Register [8], ensuring complete follow-up as only patients with Danish identification number were included and the endpoint was in-hospital mortality.

Due to the wide range of diseases admitted through our MAU, included patients were divided into major disease categories (MDC) according to the International Statistical Classification of Diseases (ICD-10) at discharge. AUROC was calculated for the specific categories to see if it can be used for a wide range of diseases.

2.3. Ethics

According to Danish law, approval by an ethics committee is not required for register-based studies. The study was reported to the Danish Data Protection Agency. The study will be reported in accordance with the TRIPOD guidelines [7].

2.4. Statistics

Data will be presented as median (interquartile range) or number (proportion) as appropriate. Differences in categorical data will be tested using Chi-squared test.

The discriminatory power of LDT-EWS was assessed by calculating the area under the receiver-operating characteristics curve (AUROC). This is the ability of a model to differentiate those with a given outcome from those without. An AUROC value > 0.8 indicates good discrimination of a model [9].

The calibration of LDT-EWS in the test population was assessed using the Hosmer-Lemeshow goodness-of-fit test. Calibration is the ability of a model to correctly estimate the risk of in-hospital death [9]. The study is reported in accordance with the TRIPOD checklist [7]. The statistical analyses were performed using Stata 14.2.