Injury, Int. J. Care Injured 46 (2015) 2325–2334

Contents lists available at ScienceDirect

Injury

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

Review

Risk scoring models for predicting peri-operative morbidity and mortality in people with fragility hip fractures: Qualitative systematic review

Takawira C. Marufu, Alexa Mannings, Iain K. Moppett *

Anaesthesia and Critical Care Research Group, Division of Clinical Neuroscience, Queen's Medical Centre, University of Nottingham, UK

ARTICLE INFO

Article history:
Accepted 13 October 2015

Keywords:
Fragility hip fracture
Risk scoring tools
Risk stratification tools
Peri-operative
Elderly

ABSTRACT

Rationale: Accurate peri-operative risk prediction is an essential element of clinical practice. Various risk stratification tools for assessing patients' risk of mortality or morbidity have been developed and applied in clinical practice over the years. This review aims to outline essential characteristics (predictive accuracy, objectivity, clinical utility) of currently available risk scoring tools for hip fracture patients.
Methods: We searched eight databases; AMED, CINHAL, Clinical Trials.gov, Cochrane, DARE, EMBASE, MEDLINE and Web of Science for all relevant studies published until April 2015. We included published English language observational studies that considered the predictive accuracy of risk stratification tools for patients with fragility hip fracture.
Results: After removal of duplicates, 15,620 studies were screened. Twenty-nine papers met the inclusion criteria, evaluating 25 risk stratification tools. Risk stratification tools considered in more than two studies were; ASA, CCI, E-PASS, NHFS and O-POSSUM. All tools were moderately accurate and validated in multiple studies; however there are some limitations to consider. The E-PASS and O-POSSUM are comprehensive but complex, and require intraoperative data making them a challenge for use on patient bedside. The ASA, CCI and NHFS are simple, easy and inexpensive using routinely available preoperative data. Contrary to the ASA and CCI which has subjective variables in addition to other limitations, the NHFS variables are all objective.
Conclusion: In the search for a simple and inexpensive, easy to calculate, objective and accurate tool, the NHFS may be the most appropriate of the currently available scores for hip fracture patients. However more studies need to be undertaken before it becomes a national hip fracture risk stratification or audit tool of choice.

Contents

* Corresponding author at: Anaesthesia and Critical Care Section, Division of Clinical Neuroscience, University of Nottingham, Queen's Medical Centre Campus, Nottingham University Hospitals NHS Trust Derby Road, Nottingham, NG7 2UH, UK. Tel.: +44 1158230959.
E-mail address: iain.moppett@nottingham.ac.uk (I.K. Moppett).

http://dx.doi.org/10.1016/j.injury.2015.10.025

Introduction

Fragility hip fractures among the elderly constitute a significant global public health problem. Risk scoring to identify high risk patients is strongly encouraged [1]. It aims to provide prognostic information based on available patient data. This in-turn allows: (a) increased objectivity in patient outcome prediction, (b) guidance on clinical decision making during perioperative period, (c) better informed consent for patients undergoing hip fracture surgery [2], and (d) treatment optimisation to improve outcome.

Various scoring tools exist and there is uncertainty as to the most suitable tool for use in hip fracture. The ideal risk scoring tool has the following attributes: simple; ease of use; reproducible; accurate; reliable; objective and available to all patients [2]. The extent to which current hip fracture scoring systems meet these criteria is unclear. This study aims to describe the components, likely clinical utility and degree of validation of published risk scoring tools.

Materials and methods

We searched eight databases; AMED, CINHAL, Clinical Trials.gov, Cochrane, DARE, EMBASE, MEDLINE, and Web of Science. The review considered all relevant published studies on development and validation of risk stratification tools in patients with fragility hip fracture. Studies were considered using the recommended standards guidelines for reporting systematic reviews of observation studies [3]. All relevant studies worldwide in any language published from 1966 to the 30th of April 2015, inclusive were included in the review. The search strategy is outlined as supplementary data appendix 1.

Study selection and outcome definition

We defined a risk stratification tool as "a scoring system or model used to predict or adjust for either mortality or morbidity after surgery, and which contains at least two different risk factors" [1]. Eligible studies were identified by title, abstract and full-text screening independently by the authors and discrepancies resolved by consensus. Manual hand searching of first generation reference lists was performed. Data extraction was independently undertaken by TM and AM on pre-piloted database forms. We extracted data for each study against the following four facets of validity and reliability: (1) development of items: development and validation samples in same or different cohorts; random selection of samples; (2) process for validation: single centre; multicentre; international; (3) metrics of discrimination: AUROC/*c*-statistics; and (4) metrics of calibration: Hosmer–Lemeshow or Pearson chi-square statistics. Studies were assessed for methodological quality and risk of bias using Altman's [4] framework for assessing internal validity.

Data and statistical analysis

Calibration and discrimination are the two main performance measures used to evaluate individual risk scoring tools. Discrimination was reported using either the AUROC or the concordance (*c*-) statistic with AUROC of less than 0.7, 0.7–0.9 and greater than 0.9 considered to indicate poor, moderate and high tool performance respectively [1]. As AUROC was not consistently reported, the observed compared to expected outcome ratio (observed/expected (O/E)), Spearman's rank correlation and chi-squared test were also used to evaluate risk scoring tool performance.

The agreement between observed and predicted outcomes (calibration) was evaluated using Hosmer–Lemeshow or Pearson chi-square statistics. $P < 0.05$ reflected evidence of lack of fit [1].

Results

The search produced 15,620 articles, and 680 were eligible for abstract screening (Fig. 1). Most studies considered at the abstract stage, reported risks for sustaining hip fracture, rather than outcome following hip fracture, and 12 studies were conference abstract presentations with no full published papers and therefore were excluded leaving 43 studies for full text analysis. Of the 43 full text studies sought, 30 [5–34] met the inclusion criteria with results presented with sufficient data to evaluate the study outcomes (Table 1). Thirteen full text studies [35–48] did not have sufficient qualitative or quantitative data relevant to this review, and were excluded. All studies included in this review were cohort studies.

Quality assessment

Quality assessment for eligible studies is outlined in Table 1. Seven studies were multicentre with a maximum of nine study sites in one study [24]. Selection bias was not observed in the included studies, though ethnic origin was constrained by the demographic of the study country. Heterogeneity among included studies was observed in method of statistical analysis, variation in time frame of outcome measurements, and in the number of models assessed by individual articles.

Validation

Three forms of validation were observed across included studies; (a) internal – validation in split sample of the same study population as tool derivation cohort, (b) external – validation in