Received: 19 December 2017 | Revised: 18 March 2018 | Accepted: 22 March 2018

DOI: 10.1111/aas.13131

ORIGINAL ARTICLE

acta Anaesthesiologica Scandinavica

Accuracy of the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity score and the Nottingham risk score in hip fracture patients in Sweden — A prospective observational study

M. H. Jonsson[1,2] | P. Bentzer[2,3] | A. Turkiewicz[2,4] | A. Hommel[2,4]

[1]Department of Anaesthesia and Intensive Care Medicine, Ystad Hospital, Ystad, Sweden

[2]Department of Clinical Sciences, Lund University, Lund, Sweden

[3]Department of Anaesthesia and Intensive Care, Helsingborg Hospital, Helsingborg, Sweden

[4]Clinical Epidemiology Unit, Orthopaedics, Lund University, Lund, Sweden

Correspondence
M. H. Jonsson, Department of Anaesthesia and Intensive Care Medicine, Ystad Hospital, Ystad, Sweden.
Email: magnushjaltalin@hotmail.com

Funding information
The study was funded by The Swedish Research Council, Region Skåne, ALF (#86626) (Swedish government funding to university hospitals for research and education) and by the Anna and Edwin Berger Foundation. No commercial funding was received.

Background: Little is known about accuracy of common risk prediction scores in elderly patients suffering from hip fractures. The objective of this study was to investigate accuracy of the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) score, Portsmouth-POSSUM (P-POSSUM) score and the Nottingham Hip Fracture Score (NHFS) for prediction of mortality and morbidity in this patient group.

Methods: This was a prospective single centre observational study on 997 patients suffering out-of-hospital cervical, trochanteric or subtrochanteric fracture of the neck of the femur. Calibration and discrimination was assessed by calculating the ratio of observed to expected events (O:E) and areas under receiver operating characteristics curves (ROC).

Results: The 30-day mortality was 6.2% and complications, as defined by POSSUM, occurred in 41% of the patients. Overall O:E ratios for POSSUM, P-POSSUM and NHFS scores for 30-day mortality were 0.90, 0.98, and 0.79 respectively. The models underestimated mortality in the lower risk bands and overestimated mortality in the higher risk bands. In contrast, POSSUM predicted morbidity well with O:E ratios close to unity in most risk bands. The areas under the ROC curves for the scoring systems was 0.60-0.67.

Conclusion: The POSSUM score and NHFS show moderate calibration and poor discrimination in this cohort. The results suggest that mortality and morbidity in hip fracture patients are largely dependent on factors that are not included in these scores.

1 | INTRODUCTION

The age standardized incidence of hip fractures in Swedish women is estimated to be around 539/100 000.[1] Similar figures have been reported for other countries in Europe and as the population is growing older the total number of hip fractures is expected to rise.[2] Patients suffering hip fractures often have multiple comorbidities and mortality and morbidity is high even though a hip fracture most commonly is a result of a low-energy trauma.[3] Treatment bundles focusing on simple measures to optimize vital organ function, pain relief and early surgical management have been introduced in several countries.[4] Even so current 30-day mortality for out-of-hospital hip fractures is 6%-10%.[3,5,6]

It is possible that outcome could be improved by identification of high-risk patients that may benefit from further work-up, additional therapeutic interventions or more advanced perioperative monitoring.[7] Furthermore, risk assessment tools may be used to compare performance within hospitals over time and between

hospitals while taking into account differences or changes in case-mix. The Physiological and Operative Severity Score for the enUmeration of mortality and morbidity (POSSUM) score is a commonly used scoring system for perioperative risks which takes into account both the preoperative physiological status and the characteristics of the surgical procedure.[8] The POSSUM score has been investigated in hip fracture patients in a number of studies with variable results with regard to calibration and discrimination.[9-14] Performance in a Scandinavian setting is unknown. The modified POSSUM equation called Portsmouth-POSSUM (P-POSSUM) was derived after reports of poor calibration of the original POSSUM score in mixed surgical cohorts. The P-POSSUM is based on the same scoring as the original POSSUM and has only been evaluated in smaller studies on hip fracture patients.[15-18]

The Nottingham Hip Fracture Score (NHFS) is specifically designed for risk prediction in elderly hip fracture patients.[19,20] It is somewhat simpler than the POSSUM and P-POSSUM scores and does not require insertion of operative data, and is thus more suited to preoperative risk evaluation. External validation has to our knowledge only been performed in United Kingdom and in the Netherlands.[12,21-23] A recent comparison between the POSSUM and the NHFS suggested that discrimination by the NHFS was superior to that of the POSSUM score.[12] Given that performance of risk scores may be influenced by differences in perioperative care between different countries and changes in care over time, it is of importance that the risk scores are evaluated and compared in new settings.

The main objective of the present study was to test the hypothesis that the POSSUM, P-POSSUM and NHFSs are well calibrated and have a high discrimination in a Swedish cohort of hip fracture patients. For this purpose, we used prospectively collected data from a single centre cohort of patients with out-of-hospital hip fractures.

2 | MATERIALS AND METHODS

2.1 | Study design

This study was based on prospectively collected data from a cohort of hip fracture patients admitted to Skåne University Hospital, Lund, Sweden between 31 January 2011 and 30 August 2014. The main purpose of the cohort was to investigate biomarkers that could predict outcome in hip fracture patients. The manuscript was prepared according to the STROBE guidelines for reporting of observational studies.[24]

2.2 | Ethics and trial registration

The study was approved by the regional ethical review board in Lund (application numbers 2010/218 and 2011/506). The study was conducted according to the Declaration of Helsinki and was monitored by external monitors from the Clinical Research Unit, Skåne University hospital, Lund. Consent was sought from patients or next of kin within 72 hours of admission. The study was registered at Clinicaltrials.gov (NCT01280253).

Editorial comment

Elderly patients with hip fractures have a high mortality risk. This study assessed the accuracy of the Possum and Nottingham risk scores for predicting 30-day mortality in a regional cohort from Sweden. Observed variations among the risk bands led the authors to suggest that there probably are other important factors which are not included in these risk models.

2.3 | Inclusion and exclusion criteria

Ambulance and emergency department staff performed patient screening. Inclusion criteria were out-of-hospital cervical, trochanteric or subtrochanteric fracture of the neck of the femur and blood sampling within 3 hours from first contact with health care providers. Patients were excluded if informed consent was not obtained, if non-operative management was chosen, if the fracture was pathological, part of multitrauma or if follow-up was not possible within the Swedish National Quality Register for Hip Fracture patients. The rationale for excluding patients with pathological fractures was that these patients were considered to represent a separate group of patients with particularly poor prognosis.[25]

2.4 | Risk scores

The POSSUM score is composed of a physiological and an operative severity score.[8] The former has 12 variables with 4 increasing values (1, 2, 4 and 8) giving a range of scores from 12 to 96. The operative score has 6 variables with the same grading and a range from 6 to 48 (Table S1). The physiological score was calculated electronically using vital parameters and laboratory values on admission, prescribed medications and comorbidities. The operative severity variable was considered as "major" as described previously.[10,18] Volume of perioperative blood loss was retrieved from anaesthesia charts. Contamination was not applicable and the variable was assigned a score of 1. Timing of operation was assigned a score of 4 for all patients ("emergency operation within 48 hours"). Outcomes were 30-day mortality and 30-day morbidity as defined in Appendix S1. Time of death was compared to admission time when defining 30-day mortality. P-POSSUM and POSSUM risk bands and calculation of predicted number of events for each risk band was done as described previously[26] and further explained in Table S3.

The NHFS is the sum of 7 variables, which are used to calculate 30-day mortality risk. The variables are age, sex, admission haemoglobin count, cognitive impairment, malignancy, independent living and number of comorbidities (Table S2). In the derivation of the score a Mini

Mental Test Score of ≤6 out of 10 was defined as cognitive impairment. Because the Short Portable Mental Status Questionnaire (SPMSQ) was established at our institution we choose to adjust the score and use a SPMSQ of ≤6 out of 10 as the definition of cognitive impairment.[27] For all other variables, definitions and cut-offs as described in the derivation of the score were used.[19] Estimation of 30-day mortality risk was performed using the modified equation from 2012.[20]

To assess if the cohort was a representative sample, demographic data and 30-day mortality data for patients fulfilling inclusion criteria on a national level during the study period were extracted from the Swedish National Quality Register for Patients with Hip Fractures.

2.5 | Statistics

No formal power analysis was performed and the number of patients included in the cohort determined the sample size. Performance of the risk scores was estimated by assessment of discrimination and calibration. Discrimination was estimated by calculating the area under the Receiving Operating Characteristics (ROC) curves for the respective scores. Areas under the ROC curve of <0.7, 0.7-0.9 and >0.9 are considered to reflect poor, moderate and high performance respectively.[28]

Calibration was evaluated by comparing observed and expected outcomes over different risk bands and for the whole cohort. Descriptive data are expressed as median and interquartile range unless stated otherwise. The accuracy of the data in the database was checked against source data by the external monitors in a sample of 31 randomly selected patients and the fraction of wrong entries was found to be <0.1%. Statistical analysis was performed using Stata™ (StataCorp. 2013. Stata Statistical Software: Release 13; StataCorp LP, College Station, TX, USA), The R Project (R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria) and SPSS (IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24; IBM Corp, Armonk, NY.)

3 | RESULTS

3.1 | Demographics

During the study period, 1845 patients were admitted to Lund University hospital with a hip fracture of which 1556 fulfilled inclusion criteria and none of the exclusion criteria. A total of 997 of these patients (64%) were included in the study. Majority of the non-included patients were screening failures in which study blood samples for the biomarker study were not collected within 3 hours. A flow chart of patients is presented in Figure 1. Demographics for included patients and eligible patients that were not included on a

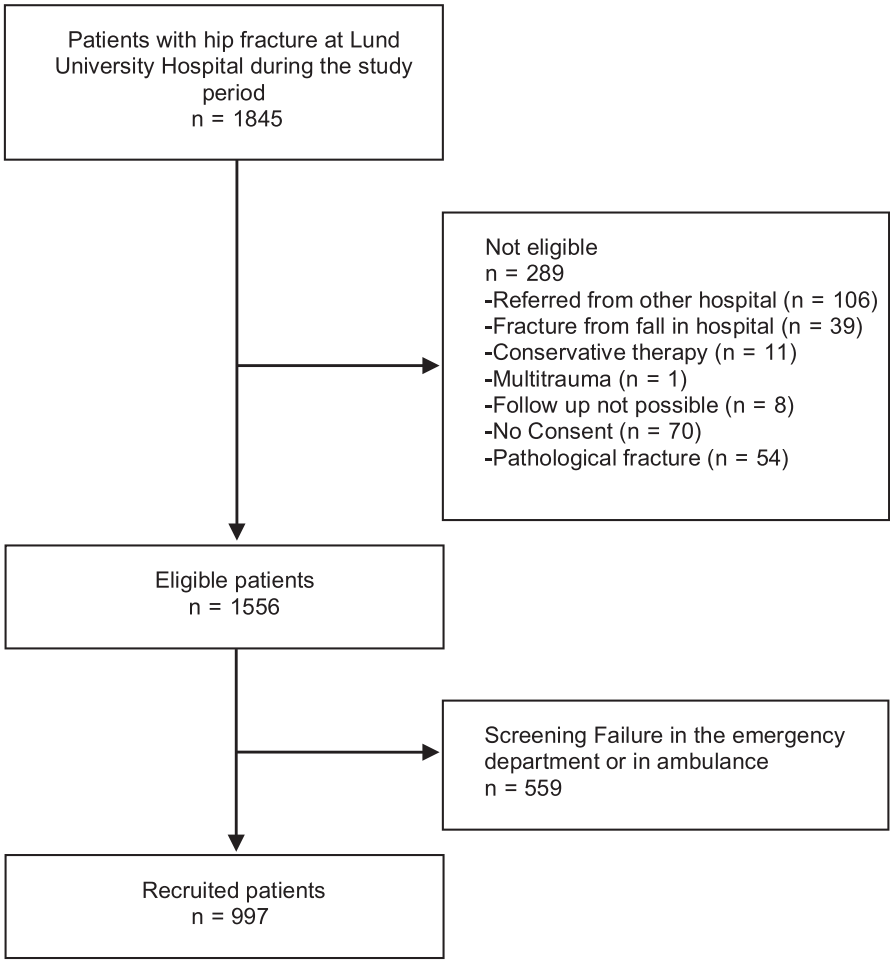

FIGURE 1 Flow chart of included patients admitted during the study period and reasons for patients not being included

national level during the study period are presented in Table 1. A detailed description of comorbidities, laboratory values and vital parameters of included patients is presented in Table 2.

3.2 | P-POSSUM mortality risk prediction

The P-POSSUM results are shown in Table 3. More than 80% of the patients had a predicted mortality risk of less than, or equal to 10% and in this risk band there were 36% more deaths observed than predicted (Table 3). All other risk bands underestimated mortality resulting in an overall prediction of 63 deaths and an O:E ratio of 0.98 (95% CI 0.76-1.21). The area under the ROC curve for prediction of 30-day mortality using the combined POSSUM physiological and operative severity score was 0.66 (95% CI 0.59-0.72) (Figure 2). Area under the ROC curve was also calculated for the physiological score separately and was 0.66 (95% CI 0.59-0.73).

TABLE 1 Demographics of the study cohort compared to patients on a national level during the study period

	Included patients	Patients fulfilling eligibility criteria on a national level[a]
Number of patients	997	48 785
Age (y)	84 (77-90)	84 (76-89)
Female (%)	71	68
ASA PS 1 (%)	5	6
ASA PS 2 (%)	29	37
ASA PS 3 (%)	62	48
ASA PS 4 (%)	4	7
Cervical fracture (%)	53	54
Trochanteric fr. (%)	41	38
Subtrochanteric fr. (%)	6	8
Time from admission to operation (h)	17 (10-23)	19 (12-25)
30-d mortality, %	6.2	7.6

ASA PS, American Society of Anesthesiologists physical status classification.
[a]Data for patients fulfilling eligibility criteria during the study period were collected from the national quality register for hip fracture patients in Sweden ("Rikshöft").

TABLE 2 Demographics of the study cohort

Patients (n)	997
Systolic blood pressure (mm Hg)	150 (130-170)
Heart rate (beats per minute)	80 (70-89)
Haemoglobin concentration (g/L)	127 (117-137)
Plasma creatinine (mmol/L)	79 (64-99)
Plasma urea (mmol/L)	7.1 (5.6-8.9)
Plasma sodium (mmol/L)	140 (138-142)
Plasma potassium (mmol/L)	3.9 (3.6-4.2)
Ischaemic heart disease, n (%)	202 (20)
Cerebrovascular disease, n (%)	208 (21)
Congestive heart failure, n (%)	133 (13)
COPD, n (%)	77 (8)
General anaesthesia, n (%)	353 (35)
Fluids administered in first 24 h (L)	3.1 (2.1-4.0)
Time from admission to operation (hrs)	18 (11-23)
More than 48 h to operation, n (%)	17 (1.7)

COPD, chronic obstructive pulmonary disease.
Continuous data are presented as median and interquartile range.

3.3 | NHFS mortality risk prediction

The NHFS results are shown in Table 4 and Figure 2. The observed mortality was 20% lower than predicted by the model. Almost 90% of the patients had a NHFS of 3-7 out of 10 and the ratio of observed to expected deaths was 0.8 for this group and 0.79 (95% CI 0.58-1.00) for the whole cohort. The area under the ROC curve for prediction of 30-day mortality was 0.67 (95% CI 0.59-0.74) (Figure 2).

3.4 | POSSUM mortality risk prediction

A total of 62 patients died within 30 days of admission, which corresponds to a mortality of 6.2%. The POSSUM mortality score predicted 69 deaths for the whole cohort resulting in an overall ratio of observed to expected (O:E ratio) deaths of 0.90 (Table S3). Almost 80% of the patients were classified as having less than a 20% risk of 30-day mortality (Table S3). In patients with a predicted risk of death of ≥20%, the POSSUM score predicted almost twice as many deaths as observed. Area under the ROC curve is identical to that of P-POSSUM as both equations use same scoring for risk prediction.

3.5 | POSSUM morbidity risk prediction

A total of 407 patients suffered at least one of the complications defined in the derivation of the POSSUM score, which corresponds to a 30-day morbidity of 41%. A detailed description of the observed

TABLE 3 Portsmouth POSSUM (P-POSSUM) mortality prediction

Risk band %	Average risk (%)	Number of patients	Expected deaths	Observed deaths	Observed: Expected death ratio
≤10	3.8	831	32	43	1.36
10-20	13.4	112	15	10	0.67
20-30	23.8	33	8	5	0.64
30-40	34.2	11	4	1	0.27
40-50	45.3	7	3	2	0.63
50-60	56.8	3	2	1	0.59
60-100	n/a	0	n/a	n/a	n/a
0 to 100	6.31	997	63	62	0.98

Risk bands are calculated using the linear method. The predicted risk of death is the average risk in each risk band. Risk bands include the upper limit but do not overlap. Overall O:E ratio 0.98 (95% CI 0.76-1.21).

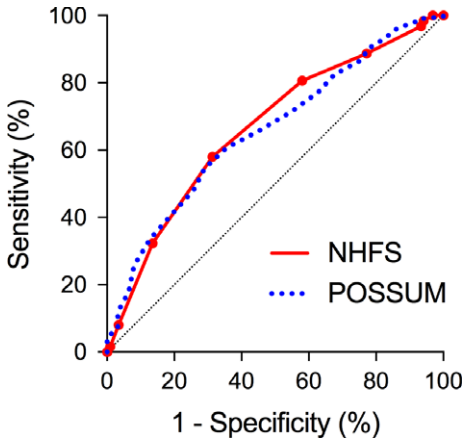

FIGURE 2 Receiver operating characteristics (ROC) curves of the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) score and Nottingham Hip Fracture Score (NHFS) for prediction of 30-d mortality. Area under the ROC curve (AUC) for POSSUM 0.66 (95% CI: 0.59-0.72); AUC for NHFS 0.67 (95% CI: 0.59-0.74) [Colour figure can be viewed at wileyonlinelibrary.com]

complications is presented in Table S5 and Appendix S1. The POSSUM morbidity score predicted a total of 411 complications resulting in an overall O:E ratio for complications of 0.99 (Table S4). In the lower (<20%) and higher risk bands (>80%) patients were few and O:E ratios indicated a poor fit (Table S4). In risk bands between 20% and 80%, POSSUM morbidity score showed O:E ratios of 0.7-1.2. The area under the ROC curve for prediction of at least 1 complication using the POSSUM morbidity score was 0.60 (95% CI 0.57-0.64).

4 | DISCUSSION

The results show that the P-POSSUM, POSSUM and NHFS scores are poorly calibrated in our cohort and underestimate mortality in the lower risk bands and overestimate mortality in the higher risk bands. In contrast, POSSUM predicted morbidity well with O:E ratios close to unity in most risk bands. Discrimination for the scores was found to be poor with areas under the ROC curves of 0.67 or less.

The original POSSUM equation was suggested to overestimate the mortality in the lower risk bands and an adjustment of the mortality equation called the P-POSSUM was later introduced.[15] However, the 2 equations were calibrated using 2 different methods and the method used in the calibration of POSSUM results in risk bands that are dependent on the characteristics of the cohort, whereas the P-POSSUM may be used to calculate risk within pre-specified risk bands (for description of the risk bands of the POSSUM mortality and morbidity see Table S3). The POSSUM score is thus limited to being an audit tool and not useful to predict outcome of individual patients. In contrast, P-POSSUM is conceptually easier to understand and allows both for prediction of outcome for individual patients and as an audit tool but has not been investigated in hip fracture patients previously. As mentioned previously, POSSUM and P-POSSUM scoring requires intraoperative data and it could be argued that this makes them less suited as screening tools. However, the operative score varies very little between the different patients and the identical areas under the ROC curves for the complete POSSUM score and the Physiological POSSUM scores, respectively, suggest that discrimination is not dependent on this parameter.

Previous studies investigating the calibration and discrimination of the POSSUM score in hip fracture patients have yielded conflicting results. Three studies with the objective to validate POSSUM in hip fracture patients have suggested that POSSUM overestimates mortality in all risk bands.[9,11,12] Other studies study reported overall O:E ratios for both mortality and morbidity that were close to unity.[10,14] Our analysis of outcome in the different risk bands revealed that calibration was poor independent of method to predict mortality. Moreover, ROC analysis in the present study suggests poor discrimination and is very similar to the area under ROC curve reported for hip fracture patients in previous studies (0.62-0.68)[9,12,13] with the exception of a single study reporting an area

TABLE 4 Nottingham Hip Fracture Score (NHFS) mortality prediction in the cohort

NHFS (points)	Number of patients	Percentage of cohort, %	Predicted 30-d mortality risk (%)	Expected deaths	Observed deaths	Observed:Expected death ratio
0	29	2.9	0.7	0	0	n/a
1	27	2.7	1.1	0	1	3.37
2	8	0.8	1.7	0	1	7.35
3	156	15.6	2.7	4	5	1.19
4	184	18.5	4.4	8	5	0.62
5	264	26.5	6.9	18	14	0.77
6	182	18.3	11	20	16	0.80
7	110	11.0	16	18	15	0.85
8	27	2.7	24	6	4	0.62
9	9	0.9	34	3	1	0.33
10	1	0.1	45	0	0	n/a
0-10	997	100	7.95	79	62	0.79

under the ROC curve of 0.83.[14] As suggested above, a potential explanation for the differing results in the studies targeting hip fracture patients is the method used to calculate risk bands. In addition, it is possible that differences in timing for collection of physiological variables and laboratory values in relation to the operation may have influenced results. Our results for hip fracture patients contrast to the good discrimination by the POSSUM score in other patient groups and suggest that parameters other than those included in the score are of importance in these elderly and fragile patients.[29,30]

Comparing complications between studies is precarious because of differences in follow-up, definitions as well as in the reliability of the registration of complications between hospitals and countries at different times. A certain amount of underestimation is likely in all cases. Furthermore, it can be argued that summarizing complications is questionable when for example the same weight is given to a large pulmonary embolus as to a urinary tract infection. Standardized outcome parameters and follow-up times are welcomed.[31] These considerations make it understandable that many previous authors have focused on the mortality prediction of the POSSUM risk score and chosen to omit the morbidity prediction. Our result that, the POSSUM morbidity prediction gave the most accurate O:E ratios over majority of risk bands in hip fracture patients agrees with a smaller study and indicate that POSSUM morbidity prediction could be useful.[16] It should, however, be noted that the area under the ROC curve was only 0.60 indicating poor discrimination.

The NHFS was developed in response to reports of poor performance of POSSUM in hip fracture patient.[9,19] Our result of an area under the ROC curve of 0.67 is within the range of 0.64-0.77 reported in previous studies[12,19,21,23,32] and suggests an equally poor discrimination for POSSUM and the NHFS in our cohort. Interestingly, while some studies have demonstrated superior discrimination by scores specifically designed for hip fracture patients[12] compared to that of general surgical risk scores several other studies have not been able to demonstrate such a difference.[19,32] Our results align with the latter studies and illustrate the importance of external validation of risk prediction scores. It should be noted that by using fewer and only preoperative variables, the NHFS is easier to use than the POSSUM score and more suited for preoperative risk assessment and may be the better alternative, as performance does not appear to differ.

4.1 | Limitations

Given that only 64% of eligible patients during the study period were included the external validity of our data could be questioned. However, our result show very similar demographics for eligible patients at our institution, and for potentially eligible patients nationally suggesting that our cohort is a representative sample and supports the external validity of our results.

While our cohort is one of the largest in which, the POSSUM score has been validated in orthopaedic surgery, both patients with high POSSUM and NHFS scores for mortality were relatively rare. This means that the validation of the scores were sensitive to single events.

5 | CONCLUSION

We conclude that the POSSUM, P-POSSUM scores and the NHFS have equally poor discrimination with regard to 30-day mortality in this cohort of elderly patients suffering hip fractures. The results suggest that mortality in this group of patients to a large extent is dependent on factors that are not included in these risk scores and highlight the need for novel approaches to identify high-risk patients.

ACKNOWLEDGEMENTS

We are deeply grateful to ambulance staff of Region Skåne, emergency ward staff and the staff at the orthopaedic wards at Skåne University hospital (Lund) for recruiting patients to this study. The following persons contributed greatly with data collection and registration: Susann Schrey, Lena Jönsson, Anne Adolfsson and Carina Lilja.

CONFLICTS OF INTEREST

The authors have no conflict of interest.

AUTHORS' CONTRIBUTIONS

M. H. J., P. B., A. T. and A. H.: Interpreted the results, critically reviewed the manuscript, and approved the final version of the manuscript for submission.

ORCID

M. H. Jonsson http://orcid.org/0000-0001-9844-8628

REFERENCES

1. Kanis JA, Oden A, McCloskey EV, et al. A systematic review of hip fracture incidence and probability of fracture worldwide. *Osteoporos Int*. 2012;23:2239-2256.
2. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet*. 2002;359:1761-1767.
3. Bergeron E, Lavoie A, Belcaid A, Ratte S, Clas D. Should patients with isolated hip fractures be included in trauma registries? *J Trauma*. 2005;58:793-797.
4. Simunovic N, Devereaux PJ, Sprague S, et al. Effect of early surgery after hip fracture on mortality and complications: systematic review and meta-analysis. *CMAJ*. 2010;182:1609-1616.
5. Boulton C. Royal College of Physicians, National Hip Fracture Database annual report 2016: Royal College of Physicians; 2016.
6. Tsang C, Boulton C, Burgon V, Johansen A, Wakeman R, Cromwell DA. Predicting 30-day mortality after hip fracture surgery: evaluation

of the National Hip Fracture Database case-mix adjustment model. *Bone Joint Res*. 2017;6:550-556.
7. Wijeysundera DN. Predicting outcomes: is there utility in risk scores? *Can J Anaesth*. 2016;63:148-158.
8. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg*. 1991;78:355-360.
9. Ramanathan TS, Moppett IK, Wenn R, Moran CG. POSSUM scoring for patients with fractured neck of femur. *Br J Anaesth*. 2005;94:430-433.
10. Wright DM, Blanckley S, Stewart GJ, Copeland GP. The use of orthopaedic POSSUM as an audit tool for fractured neck of femur. *Injury*. 2008;39:430-435.
11. Young W, Seigne R, Bright S, Gardner M. Audit of morbidity and mortality following neck of femur fracture using the POSSUM scoring system. *N Z Med J*. 2006;119:U1986.
12. Karres J, Heesakkers NA, Ultee JM, Vrouenraets BC. Predicting 30-day mortality following hip fracture surgery: evaluation of six risk prediction models. *Injury*. 2015;46:371-377.
13. Burgos E, Gomez-Arnau JI, Diez R, Munoz L, Fernandez-Guisasola J, Garcia del Valle S. Predictive value of six risk scores for outcome after surgical repair of hip fracture in elderly patients. *Acta Anaesthesiol Scand*. 2008;52:125-131.
14. van Zeeland ML, Genovesi IP, Mulder JW, Strating PR, Glas AS, Engel AF. POSSUM predicts hospital mortality and long-term survival in patients with hip fractures. *J Trauma*. 2011;70:E67-E72.
15. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. *Br J Surg*. 1998;85:1217-1220.
16. Bonicoli E, Parchi P, Piolanti N, Andreani L, Niccolai F, Lisanti M. Comparison of the POSSUM score and P-POSSUM score in patients with femoral neck fracture. *Musculoskelet Surg*. 2014;98:201-204.
17. Hirose J, Ide J, Irie H, Kikukawa K, Mizuta H. New equations for predicting postoperative risk in patients with hip fracture. *Clin Orthop Relat Res*. 2009;467:3327-3333.
18. Mohamed K, Copeland GP, Boot DA, et al. An assessment of the POSSUM system in orthopaedic surgery. *J Bone Joint Surg Br*. 2002;84:735-739.
19. Maxwell MJ, Moran CG, Moppett IK. Development and validation of a preoperative scoring system to predict 30 day mortality in patients undergoing hip fracture surgery. *Br J Anaesth*. 2008;101:511-517.
20. Moppett IK, Parker M, Griffiths R, Bowers T, White SM, Moran CG. Nottingham Hip Fracture Score: longitudinal and multi-assessment. *Br J Anaesth*. 2012;109:546-550.
21. Marufu TC, White SM, Griffiths R, Moonesinghe SR, Moppett IK. Prediction of 30-day mortality after hip fracture surgery by the Nottingham Hip Fracture Score and the Surgical Outcome Risk Tool. *Anaesthesia*. 2016;71:515-521.
22. Rushton PR, Reed MR, Pratt RK. Independent validation of the Nottingham Hip Fracture Score and identification of regional variation in patient risk within England. *Bone Joint J*. 2015;97-B:100-103.
23. Nijmeijer WS, Folbert EC, Vermeer M, Slaets JP, Hegeman JH. Prediction of early mortality following hip fracture surgery in frail elderly: the Almelo Hip Fracture Score (AHFS). *Injury*. 2016;47:2138-2143.
24. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg*. 2014;12:1500-1502.
25. Sarahrudi K, Greitbauer M, Platzer P, Hausmann JT, Heinz T, Vecsei V. Surgical treatment of metastatic fractures of the femur: a retrospective analysis of 142 patients. *J Trauma*. 2009;66:1158-1163.
26. Wijesinghe LD, Mahmood T, Scott DJ, Berridge DC, Kent PJ, Kester RC. Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. *Br J Surg*. 1998;85:209-212.
27. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Geriatr Soc*. 1975;23:433-441.
28. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MP. Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: qualitative systematic review. *Anesthesiology*. 2013;119:959-981.
29. Igari K, Ochiai T, Yamazaki S. POSSUM and P-POSSUM for risk assessment in general surgery in the elderly. *Hepatogastroenterology*. 2013;60:1320-1327.
30. Richards CH, Leitch FE, Horgan PG, McMillan DC. A systematic review of POSSUM and its related models as predictors of postoperative mortality and morbidity in patients undergoing surgery for colorectal cancer. *J Gastrointest Surg*. 2010;14:1511-1520.
31. Jammer I, Wickboldt N, Sander M, et al. Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions: a statement from the ESA-ESICM joint taskforce on perioperative outcome measures. *Eur J Anaesthesiol*. 2015;32:88-105.
32. Boddaert J, Na N, Le Manach Y, et al. Prediction of postoperative mortality in elderly patients with hip fracture: are specific and geriatric scores better than general scores? *Br J Anaesth*. 2017;118:952-954.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Jonsson MH, Bentzer P, Turkiewicz A, Hommel A. Accuracy of the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity score and the Nottingham risk score in hip fracture patients in Sweden—A prospective observational study. *Acta Anaesthesiol Scand*. 2018;62:1057–1063. https://doi.org/10.1111/aas.13131