

CRITICAL CARE

Development and validation of a preoperative scoring system to predict 30 day mortality in patients undergoing hip fracture surgery

M. J. Maxwell¹, C. G. Moran² and I. K. Moppett^{1*}

¹Department of Anaesthesia and ²Department of Trauma and Orthopaedics, University of Nottingham, Queen's Medical Centre Campus, Nottingham University Hospitals NHS Trust, Nottingham NG7 2UH, UK

*Corresponding author. E-mail: iain.moppett@nottingham.ac.uk

Background. Hip fractures are common in the elderly and have a high 30 day postoperative mortality. The ability to recognize patients at high risk of poor outcomes before operation would be an important clinical advance. This study has determined key prognostic factors predicting 30 day mortality in a hip fracture population, and incorporated them into a scoring system to be used on admission.

Methods. A cohort study was conducted at the Queen's Medical Centre, Nottingham, over a period of 7 yr. Complete data were collected from 4967 patients and analysed. Forward univariate logistic regression was used to select the independent predictor variables of 30 day mortality, and then multivariate logistic regression was applied to the data to construct and validate the scoring system.

Results. The variables found to be independent predictors of mortality at 30 days were: age (66–85 yr, ≥ 86 yr), sex (male), number of co-morbidities (≥ 2), mini-mental test score (≤ 6 out of 10), admission haemoglobin concentration (≤ 10 g dl⁻¹), living in an institution, and presence of malignant disease. These variables were subsequently incorporated into a risk score, the Nottingham Hip Fracture Score. The number of deaths observed at 30 days, and the number of deaths predicted by the scoring system, indicated good concordance (χ^2 test, $P=0.79$). The area (SE) under the receiver operating characteristic curve was 0.719 (0.018), which demonstrated a reasonable predictive value for the score.

Conclusions. We have developed and validated a scoring system that reliably predicts the probability of mortality at 30 days for patients after hip fracture.

Br J Anaesth 2008; **101**: 511–17

Keywords: complications, death; complications, trauma; risk; surgery, orthopaedic

Accepted for publication: June 15, 2008

Hip fractures are frequent in elderly individuals and their incidence is rising. More than 80 000 patients in the UK sustain fractures of their proximal femur annually.¹ As elderly people remain the fastest growing section of our society, this number will continue to climb, in spite of attempts at primary and secondary fracture prevention. Hip fractures create an enormous burden, both clinical and economic, on the health service. They can also have a devastating impact on patients and their families, with a high 30 day postoperative mortality and many patients unable to return home.

Much work has been published focusing on the factors which influence patients' outcome after hip fracture, and numerous factors have been shown to affect morbidity and mortality. To date, however, these factors have been of limited utility in clinical practice. Preoperative recognition of patients at particularly high risk for adverse outcome may be useful for several reasons: appropriate informed consent; timing of surgery; access to higher level care before or after operation; and intra- and interdepartmental audit.

Scoring systems to predict outcome are available. Although they work well for a general population (e.g.

physiological and operative severity score for enumeration of mortality and morbidity, POSSUM),² they are of limited validity in the specific, relatively homogeneous population of hip fracture patients.³

Since 1999, the Queen's Medical Centre in Nottingham has collected audit data on all patients admitted with a fractured neck of femur. We have used this information to determine which prognostic factors are most important at predicting 30 day mortality in our population, and developed a scoring system, the Nottingham Hip Fracture Score (NHFS), to help identify these patients on admission.

Methods

From May 1999 to April 2006, all patients admitted with a fractured neck of femur to the Queen's Medical Centre have had prospective collection of physiological and operative data with the intention of auditing factors associated with this population's morbidity and mortality. Data are collected prospectively by dedicated audit officers from hospital computer and paper records, and from the patients themselves. Data from patients who do not undergo surgery are included in the data collection. Thirty-day mortality data are collected and cross-checked both from hospital statistics and from the Office of National Statistics. Individual physiological values which deviated noticeably from normal, or any missing data points, were cross-checked manually with hospital records by one of the investigative team. The accuracy of the data has previously been verified, with internal cross-checking demonstrating an error rate of <3%. Database management ensured patient anonymity and confidentiality, and fully complied with the Caldicott principles.

Selection of predictor variables

All data were entered into an SPSS for Windows[®] spreadsheet (Version 14.0). Univariate logistic regression analysis was then performed on all potential variables in order to select those which were predictors of mortality at 30 days. The variables considered were based upon factors highlighted as significant in previously published research. The included variables were: age,^{4–11} sex,^{4–6 8 10–16} admission mini-mental test score,^{5 11 15–17} admission haemoglobin^{18 19} and urea²⁰ concentrations, presence on admission of cardiovascular,^{12 21 22} cerebrovascular,²³ respiratory,^{21 24} or malignant^{7 12 24} diseases, independence of basic activities of daily living,^{5 7} living in an institution,^{4 22} and total number of co-morbidities (cardiovascular, cerebrovascular, respiratory, renal diseases, diabetes, and systemic malignancy). Pre-existing diseases were defined on the basis of the admission history from the patient, relatives, or notes as: cardiovascular disease (pre-existing cardiovascular conditions including previous myocardial infarction, angina, atrial fibrillation, valvular heart disease, or hypertension); cerebrovascular disease (patient has suffered a stroke or transient ischaemic attack in their

lifetime); respiratory disease (pre-existing chronic respiratory conditions, including asthma or chronic obstructive airways disease but not including acute infections); renal disease (pre-existing known renal disease but not elevated urea without diagnosis of a renal condition); and malignancy (active malignancy within 20 yr but not including non-invasive skin cancer). Data on patients who did not undergo surgery are reported for comparison with the data used to derive the hip fracture score.

Development and validation of the scoring system

Each patient in the database was assigned at random to a creation or validation data set. The former group was used to develop the scoring system, and the latter to evaluate the score's performance and accuracy. An automated step-wise forward multivariate logistic regression analysis was then applied to the data within the creation set in order to construct the score. Variables identified from the univariate analysis as potential predictors were included in the multivariate analysis. The *P*-value for entry into the model was 0.05 and for removal 0.1. Indicator variables were used for analysis of categorical variables and were selected to keep all odds ratios greater than unity. Mortality at 30 days after fracture was entered as the dependent variable and the independent predictor variables were entered as covariates. Interaction between covariates was tested for formally. The final coefficients produced were then altered by rounding to integer values to create a more easily applicable, and clinically usable, scoring system.

Age as a covariate was examined in two ways. First, as a categorical variable split at ≤ 65 , 66–85, and ≥ 86 yr, and secondly as a linear variable, censoring age at 0, 65, or 85 yr. In other words, individuals at or below the censored age were assigned an age of 0 yr and individuals above that age were given an adjusted age of 'real age–censored age'. A similar process was used for haemoglobin concentration. Analyses were performed using a categorical variable with a cut-off 10 g dl⁻¹ or as a simple linear variable. The number of co-morbidities was analysed in two ways: as an ordinal variable from 0 to 5 or as a categorical variable split between 0–1 and ≥ 2 co-morbidities.

The performance of the score was assessed using the data within the validation set. Goodness-of-fit of the score to the data was assessed using the Hosmer–Lemeshow statistic.²⁵ This statistic divides cases into 10 approximately equal-sized groups, according to increasing score values, and compares the predicted with the observed death rates. A lack of difference between the predicted and the observed deaths indicates good concordance of the score. Sensitivity analysis was performed using standard receiver operating characteristic (ROC) curves. For comparison with potentially simpler scoring systems, two other systems were analysed. The American Society of Anesthesiology (ASA) classification²⁶ was used as a risk predictor, generating risk bands using the same creation data set. The Donati score²⁷

was also calculated. This is a four-item score using age, ASA, surgical severity, and urgency of surgery. The latter two items are essentially fixed for this population, so the score assesses the effect of age and ASA on risk.

Results

Over the 7 yr of the study, 5162 patients were admitted with fractured neck of femur to the Queen’s Medical Centre. Four thousand nine hundred and sixty-seven of these patients underwent operative management and have therefore been included in the development of the scoring system.

Two thousand four hundred and ninety-two patients (50.2%) were randomly assigned to the creation group, and the remaining 2475 patients (49.8%) were assigned to the validation group. The characteristics of the two groups and the characteristics of the patients who were not operated on are summarized in Tables 1 and 2. The variables analysed as predictors of 30 day mortality on univariate analysis are shown in Table 3. Those variables which were significant predictors (i.e. for which the 95% confidence limit of the odds ratio did not include unity) were then included in the multivariate analysis.

Score system development

The following variables were found to be independent predictors of 30 day mortality on multivariate logistic regression analysis: age (66–85 and ≥86), sex (male), number of co-morbidities (≥2), mini-mental test score (≤6 out of 10), admission haemoglobin concentration (≤10 g dl⁻¹), living in an institution, and presence of malignant disease.

Table 1 Characteristics of patients in the creation and validation groups. sd, standard deviation; Hb, haemoglobin; MMTS, mini-mental test score

	Creation group	Validation group
Number of cases	2492	2475
Age [mean (range)]	79.9 (17–102)	80.1 (17–101)
Sex (n, % males)	587 (23.6)	566 (22.9)
Admission Hb mg dl ⁻¹ [mean (sd)]	12.3 (1.8)	12.3 (1.8)
Admission MMTS ≤6 out of 10 (n, %)	922 (40.0)	847 (34.2)
Living in an institution (n, %)	645 (25.9)	599 (24.2)
Patients with two or more co-morbidities (n, %)	595 (23.9)	560 (22.6)
Death within 30 days (n, %)	200 (8.0)	190 (7.7)

Table 2 Characteristics of patients who did not undergo surgery. sd, standard deviation; Hb, haemoglobin; MMTS, mini-mental test score

Number of cases	195
Age [mean (range)]	79.5 (17–102)
Sex (n, % males)	76 (39.0)
Admission Hb mg dl ⁻¹ [mean (sd)]	11.7 (2.0)
Admission MMTS ≤6 out of 10 (n, %)	88 (45.1)
Living in an institution (n, %)	76 (40.0)
Patients with two or more co-morbidities (n, %)	84 (43.1)
Presence of malignant disease (n, %)	36 (18.5)
Death within 30 days (n, %)	74 (37.9)

Using age as a linear variable (either from 0, 65, or 85 yr) worsened the goodness-of-fit of the model, and the area under the ROC curve remained the same. Similarly, treating haemoglobin as a linear variable also led to a worse goodness-of-fit and a similar area under the curve (AUC). There was a small interaction effect of age and number of co-morbidities; however, the change in odds ratio was small, and again there was no improvement in model fit if the interaction was included. In order to keep the score relatively simple, these effects were left out of the final model.

The NHFS was then produced by multiplying the coefficients by two and rounding them to the nearest integer value. The score ranges from 0 to 10 points. The results of the multivariate logistic regression analysis are presented in Table 4.

To calculate the percentage of patients predicted to have died at our unit within 30 days after fractured neck of femur, the following logistic equation is used:

$$30 \text{ day mortality (\%)} = \frac{100}{1 + e^{(4.718 - (\text{NHFS}/2))}}$$

This equation was generated by the final step of the logistic regression process and requires the patient’s total NHFS to be entered.

Table 3 Results of univariate logistic regression analysis. MMTS, mini-mental test score; Hb, haemoglobin; ADLs, activities of daily living

Factor	Value	Odds ratio	95% confidence interval
Age	66–85 yr	3.68	1.99–6.79
	≥86 yr	6.66	3.6–12.3
Sex	Male	1.69	1.38–2.07
	MMTS	≤6 out of 10	2.42
Co-morbidities (0–1 vs ≥2)	≥2	2.10	1.72–2.56
Cardiovascular disease	Yes	1.40	1.16–1.70
Respiratory disease	Yes	2.01	1.61–2.50
Cerebrovascular disease	Yes	1.34	1.04–1.73
Malignant disease	Yes	1.94	1.50–2.50
Admission Hb	≤10 g dl ⁻¹	1.87	1.45–2.42
Admission urea	≥12 mmol litre ⁻¹	1.97	1.58–2.46
Living in an institution	Yes	2.23	1.83–2.7
Independence of basic ADLs	No	0.95	0.77–1.19
Pre-fracture mobility	With aids	1.35	1.12–1.65

Table 4 Results of multivariate logistic regression analysis. Hb, haemoglobin; MMTS, mini-mental test score

Variable	Value	Coefficient	Odds ratio	95% confidence interval	NHFS Score
Age	66–85 yr	1.468	4.34	1.34–14.0	3
	≥86 yr	1.986	7.28	2.22–23.90	4
Sex	Male	0.505	1.66	1.15–2.39	1
	Admission Hb	≤10 g dl ⁻¹	0.441	1.55	1.01–2.39
MMTS	≤6 out of 10	0.456	1.577	1.10–2.27	1
Living in an institution	Yes	0.411	1.508	0.976–2.33	1
Number of co-morbidities	≥2	0.490	1.63	1.15–2.32	1
Malignancy	Yes	0.564	1.76	1.13–2.74	1
Constant		–4.721			

Worked example

An example of NHFS usage in everyday clinical practice is given in Figure 1. Calculations of the predicted probability of 30 day mortality associated with every possible NHFS points total were then made. These results are displayed in Figure 2 and help to quantify the magnitude of risk associated with scores into clinically meaningful indices. By reference to this graph, every patient can have their predicted 30 day mortality easily calculated on admission.

Scoring system validation

The validation set data were then used to assess the performance of the score. Calculation of the Hosmer–Lemeshow goodness-of-fit statistic for each decile of risk showed that there was good concordance between observed and predicted deaths at 30 days (χ^2 test, $P=0.79$). In contrast, the Donati score showed poor concordance with $P<0.01$. It was not possible to perform a meaningful Hosmer–Lemeshow test on the ASA score. Using an NHFS ≥ 5 to predict the probability of 30 day mortality at $>10\%$ produces a sensitivity of 44.2% and a specificity of 80.8%.

- Female
- 77 yr
- Admission MMTS 5/10
- Lives in nursing home
- COPD
- Previous CVA
- Chronic renal failure
- Angina
- Admission Hb 9.4 g dl^{-1}

NOTTINGHAM HIP FRACTURE SCORE		
Variable	Value	Points
Age	66–85 yr	3
	≥ 86 yr	0
Sex	Male	0
Admission Hb	$\leq 10 \text{ g dl}^{-1}$	1
Admission MMTS	≤ 6 out of 10	1
Living in an institution	Yes	1
Number of co-morbidities	≥ 2	1
Malignancy	Yes	0
TOTAL		7

$$\begin{aligned}
 \text{PREDICTED 30 DAY MORTALITY} &= 100 / [1 + e^{(4.718 - (\text{NHFS}/2))}] \\
 &= 100 / [1 + e^{(4.718 - (7/2))}] \\
 &= 23\%
 \end{aligned}$$

Fig 1 A clinical example of the Nottingham Fracture Score. The individual clinical factors are scored, and then the risk of death at 30 days can be calculated.

The ROC curve is shown in Figure 3. A perfect test would produce a right angle with an AUC of 1. The AUC for the NHFS is 0.719 (SE 0.0184), demonstrating a reasonable predictive value for the score. ASA and Donati scores produced very similar AUC [0.718 (0.0163) and 0.717 (0.0184)].

Discussion

Using prospectively gathered data, we have developed and validated a novel risk scoring system to predict 30 day postoperative mortality in patients sustaining a fractured neck of femur. Previous workers have demonstrated various factors which associate with an increased risk of death after fractured neck of femur, but we believe we are the first to have assessed these factors together.

For a scoring system to be useful clinically, it has to fulfil several criteria: it should use readily available and verifiable clinical information; it should have been developed and validated in the population in whom it is to be used; and it should be free from confounding factors. The NHFS uses data that are easily collectable for any patient presenting

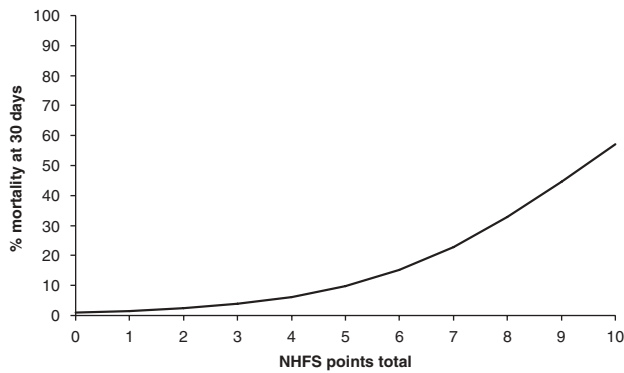


Fig 2 Graphical representation of predicted mortality at 30 days after fractured neck of femur using the NHFS.

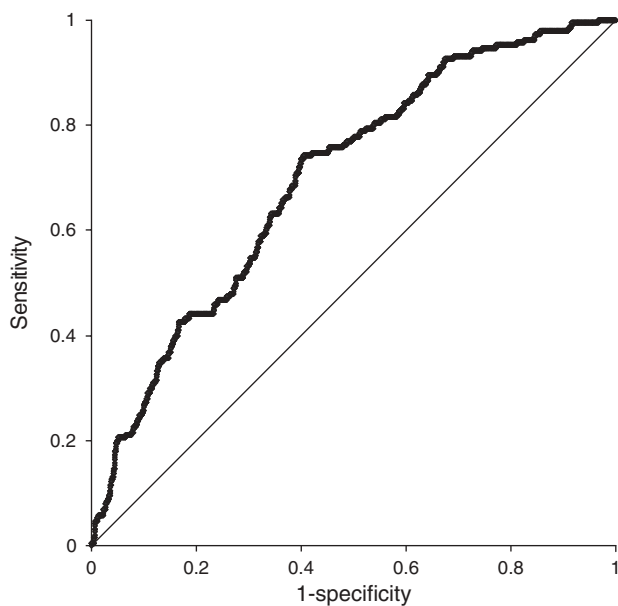


Fig 3 Validation data set ROC curve for 30 day mortality after fractured neck of femur using the NHFS.

with a fractured neck of femur. Unlike POSSUM,² we have deliberately excluded surgical and anaesthetic data. This is because we believe it is possible for surgical and anaesthetic techniques to be influenced by preoperative patient factors to unknown and varying degrees. For example, some anaesthetists will always use spinal anaesthesia for the sicker patients, whereas others prefer general anaesthesia; similarly, some surgeons will vary the form of femoral fixation depending on the preoperative state of the patient.

The concept of summing the number of diseases present is not a new one, and we realize that the use of a semi-quantitative value of ‘number of co-morbidities’ may appear controversial to some. We believe, however, that this approach should not be regarded as any more or less rigorous as considering whether specific diseases are present or absent. The universally used ASA classification makes no attempt to define the nature of the diseases which are causing limitation or threat to health. It is also

important to remember that arbitrary cut-offs always exist when defining whether a particular disease is significant or not. The information on number of co-morbidities in this study was based on the admission clerking, so it therefore reflected what the admitting orthopaedic doctor felt were important ongoing medical problems. The number of co-morbidities is a semi-quantitative substitute for the anaesthetist’s ‘end of the bed’ feeling before surgery.

Residence before hospital admission is a surrogate for general fitness, and may be affected by various factors unrelated to eventual outcome, such as proximity and ability of family carers. Often tests such as nutritional screening and ability to perform specific tasks may appear more rigorous; however, two problems exist with performing these measurements. First, they are largely impractical to implement in routine practice, and secondly, documentation of pre-morbid activity only occurs after the fracture has occurred, when recall and other biases tend to make such measures imprecise.

Validation of scoring systems is always a concern and we have attempted to address this in this paper. ROC curves are often presented, as here, although in the context of population risk of death, we would question how useful they are. Each point on the ROC curve represents a score above which all patients are predicted to die and below which all are predicted to live. The perfect test cut-off would have perfect sensitivity (ruling out all of those who will survive) and perfect specificity (ruling in only those who are to die). In clinical practice, we accept that we are trying to predict risk (a probability) not death (an event) for any one individual. The area under the ROC curve of 0.719 is reasonable, and better than orthopaedic POSSUM when applied to this population.³ It is not as good, however, as POSSUM² or the Donati score²⁷ for a general surgical or orthopaedic population, and there are probably several reasons for this. First, the aforementioned scores both include a large subset of low-risk patients: those who are young, healthy, and undergoing minor surgery. Consequently, very high sensitivity can be obtained for this subset. In contrast, patients with a fractured neck of femur represent the other end of the spectrum; the surgical insult is significant and the majority of the patients are elderly with some medical problems before admission. There is no truly low-risk population sub-set. Secondly, the significance of prognostic indicators for patients presenting with fractured neck of femur is also different, which results in a reduction in the discriminatory power of any score. A young patient presenting for surgery with renal impairment and two co-morbidities would be viewed as unusual and probably at higher risk than others. In the fractured neck of femur population, such a presentation is not unusual. POSSUM is the sum of two scores, physiological and surgical. The surgical score is the same for almost every patient in the cohort, thereby reducing the possible range of scores considerably, which inevitably reduces the power of a scoring system. The Donati score

uses age, ASA score, urgency, and operative severity as its indices.²⁷ Again, the range of scores then becomes limited, since all surgery is urgent, of the same severity and the median patient age is more than 80. Of note, the cohort used for the derivation of the Donati score varies significantly from that of the fractured neck of femur patient: only 3% were urgent, 15% were ASA III, 50% male, and the mean age was 55. The validity and applicability of such a system to the specific fractured neck of femur population is therefore questionable. The ASA score alone is poor as a discriminator in this population, since around 50% of patients are scored as ASA III, with 30 day mortality in this group of 10%, the same as for the entire cohort. ASA score therefore adds little prognostic information beyond that given by a diagnosis of hip fracture. Furthermore, ASA rating has been shown to be unreliable, particularly with regards to ASA II and III.²⁶

Hip fracture surgery is non-elective and clinicians have to balance the risk of operating on the unfit patient against the risk of delaying surgery for investigation or treatment. Goodness-of-fit tests are more relevant evaluations for scoring systems which are concerned with risk. They assess how well the score predicts outcome for bands of risk, and in the case of the Hosmer–Lemeshow test, these bands are deciles. In both the creation and validation data sets, we have demonstrated good concordance between predicted and observed outcomes, and this suggests that the score performs well across the spectrum of risk. We have found that the Donati score performs poorly in this regard and we have previously shown similar poor concordance for POSSUM.³ A recent paper from a Spanish group provides some support for our results.²⁸ They assessed the predictive value of six essentially objective scoring systems (ASA, Goldman, Charlson, POSSUM, Barthel, and ‘RISK-VAS’—an assessment of risk by an experienced observer). They found that none of the systems predicted outcome particularly well, with ASA having an area under the ROC curve not significantly different from chance.

In order to be useful, clinicians need to have confidence that a scoring system will be applicable to different centres caring for similar populations. Again, this is a matter of opinion since by definition the scoring system only applies directly to the original population. Thirty-day mortality in this series is 10.2%, and so is in line with national statistics and other published data.¹³ Patients are operated on in dedicated trauma theatres by consultants or senior trainees with an interest in trauma. Anaesthesia is provided by consultants or senior trainees. Hence, we feel our practice is a valid representation of normal UK practice. We cannot exclude, however, that aspects of our practice may differ from other units. Future work will address the use of this scoring system in other units.

Ultimately, what use is a preoperative scoring system? We suggest it will be of value in several areas. An important component of modern medical practice is internal audit; validated scoring systems allow units to assess their

performance adjusted for risk. Outcome comparisons (league tables) are also increasingly common and risk adjustment is an important part of that approach. The national hip fracture database is shortly to be launched by the British Orthopaedic Association and British Geriatric Society. It is based upon the Myocardial Infarction National Audit Project (MINAP) and the Scottish Hip Fracture Audit, and aims to collect data on hip fracture management from all hospitals in the UK. A validated scoring system will allow adjustment for variables and allow more accurate comparison of mortality between units. Although ‘end of the bed’ assessments such as ASA or the Donati score may have similar ROC curves to the NHFS, we have found that they do not fit the entire data set well. Furthermore, ASA is a subjective and unverifiable datum. The data used by the NHFS is generally collected as a matter of routine in medical and nursing notes and can be recorded and verified by clerical staff, which facilitates its use as a scoring system.

Accurate assessment of risk may affect how anaesthetists and surgeons manage their patients. It allows informed discussions with patients and relatives about likely outcome, and enables rational decisions to be made regarding the use of higher level care facilities. Clinical management may also change on the basis of predicted risk. Invasive monitoring may be withheld from those at low risk and more aggressive measures may be instituted for those at high risk, such as cardiac output monitoring or postoperative care in high level care areas. As yet we have no evidence that such measures would change the outcome in specific risk groups, but risk stratification may be a useful tool for future clinical trials. Trials can be designed to include subjects where significant changes in outcome are possible, such as those with intermediate or high levels of risk.

In conclusion, we have developed and validated a novel multi-factorial scoring system to predict risk in patients undergoing surgery for fractured neck of femur. We plan to incorporate this scoring system and predicted mortality graph into our hospital’s fractured neck of femur clerking proforma with the intention of calculating every patient’s mortality risk on admission. We believe the NHFS will be of use for both clinicians and researchers in this expanding, high-risk population.

Acknowledgements

We thank Russell Wenn, Chris Boulton, and Chris White for help with data collection. We are grateful to the consultant trauma surgeons at the Queen’s Medical Centre, Nottingham, for their support and for allowing their patients to participate in this study.

References

- 1 The Information Centre. Hospital episode statistics. Available from <http://www.hesonline.nhs.uk> (accessed November 2, 2007)
- 2 Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg* 1991; **78**: 355–60

- 3 Ramanathan TS, Moppett IK, Wenn R, Moran CG. POSSUM scoring for patients with fractured neck of femur. *Br J Anaesth* 2005; **94**: 430–3
- 4 Lu-Yao GL, Baron JA, Barrett JA, Fisher ES. Treatment and survival among elderly Americans with hip fractures: a population-based study. *Am J Public Health* 1994; **84**: 1287–91
- 5 Elliott J, Beringer T, Kee F, Marsh D, Willis C, Stevenson M. Predicting survival after treatment for fracture of the proximal femur and the effect of delays to surgery. *J Clin Epidemiol* 2003; **56**: 788–95
- 6 Walker N, Norton R, Vander Hoorn S, et al. Mortality after hip fracture: regional variations in New Zealand. *N Z Med J* 1999; **112**: 269–71
- 7 Aharonoff GB, Koval KJ, Skovron ML, Zuckerman JD. Hip fractures in the elderly: predictors of one year mortality. *J Orthop Trauma* 1997; **11**: 162–5
- 8 Roberts SE, Goldacre MJ. Time trends and demography of mortality after fractured neck of femur in an English population, 1968–98: database study. *Br Med J* 2003; **327**: 771–5
- 9 Shah MR, Aharonoff GB, Wolinsky P, Zuckerman JD, Koval KJ. Outcome after hip fracture in individuals ninety years of age and older. 2001. *J Orthop Trauma* 2003; **17**: S6–11
- 10 Fox HJ, Pooler J, Prothero D, Bannister GC. Factors affecting the outcome after proximal femoral fractures. *Injury* 1994; **25**: 297–300
- 11 Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Survival experience of aged hip fracture patients. *Am J Public Health* 1989; **79**: 274–8
- 12 Myers AH, Robinson EG, Van Natta ML, Michelson JD, Collins K, Baker SP. Hip fractures among the elderly: factors associated with in-hospital mortality. *Am J Epidemiol* 1991; **134**: 1128–37
- 13 Wehren LE, Hawkes WG, Orwig DL, Hebel JR, Zimmerman SI, Magaziner J. Gender differences in mortality after hip fracture: the role of infection. *J Bone Miner Res* 2003; **18**: 2231–7
- 14 Endo Y, Aharonoff GB, Zuckerman JD, Egol KA, Koval KJ. Gender differences in patients with hip fracture: a greater risk of morbidity and mortality in men. *J Orthop Trauma* 2005; **19**: 29–35
- 15 Muraki S, Yamamoto S, Ishibashi H, Nakamura K. Factors associated with mortality following hip fracture in Japan. *J Bone Miner Metab* 2006; **24**: 100–4
- 16 Nightingale S, Holmes J, Mason J, House A. Psychiatric illness and mortality after hip fracture. *Lancet* 2001; **357**: 1264–5
- 17 Marottoli RA, Berkman LF, Leo-Summers L, Cooney LM Jr. Predictors of mortality and institutionalization after hip fracture: the New Haven EPESE cohort. Established Populations for Epidemiologic Studies of the Elderly. *Am J Public Health* 1994; **84**: 1807–12
- 18 Gruson KI, Aharonoff GB, Egol KA, Zuckerman JD, Koval KJ. The relationship between admission hemoglobin level and outcome after hip fracture. *J Orthop Trauma* 2002; **16**: 39–44
- 19 Halm EA, Wang JJ, Boockvar K, et al. The effect of perioperative anemia on clinical and functional outcomes in patients with hip fracture. *J Orthop Trauma* 2004; **18**: 369–74
- 20 Kuokkanen HO, Korkala OL. Factors affecting survival of patients with hip fractures. *Acta Orthop Belg* 1992; **58**: 425–8
- 21 Nettleman MD, Alsip J, Schrader M, Schulte M. Predictors of mortality after acute hip fracture. *J Gen Intern Med* 1996; **11**: 765–7
- 22 Eiskjaer S, Ostgard SE. Risk factors influencing mortality after bipolar hemiarthroplasty in the treatment of fracture of the femoral neck. *Clin Orthop Relat Res* 1991; **295**: 295–300
- 23 Ramnemark A, Nilsson M, Borssen B, Gustafson Y. Stroke, a major and increasing risk factor for femoral neck fracture. *Stroke* 2000; **31**: 1572–7
- 24 Roche JJ, Wenn RT, Sahota O, Moran CG. Effect of comorbidities and postoperative complications on mortality after hip fracture in elderly people: prospective observational cohort study. *Br Med J* 2005; **331**: 1374
- 25 Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; **115**: 92–106
- 26 Owens WD, Felts JA, Spitznagel EL Jr. ASA physical status classifications: a study of consistency of ratings. *Anesthesiology* 1978; **49**: 239–43
- 27 Donati A, Ruzzi M, Adrario E, et al. A new and feasible model for predicting operative risk. *Br J Anaesth* 2004; **93**: 393–9
- 28 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**: 125–31