Validation of the Nottingham Hip Fracture Score as a predictor of 30-day mortality after hip surgery

Marcelo B. Mathias,[1] Isabela de M. Rosa,[1] Pedro José Labronici,[1] Vinicius S. Gameiro,[1] Eduardo B. de Sousa,[1] Rodrigo S. Pires e Albuquerque[1*]

Abstract

Purpose: To assess whether the Nottingham Hip Fracture Score (NHFS) can predict mortality in the first 30 days after hip surgery. Material and methods: Upon admission, 216 patients were assessed for age, sex, mobility status (bedridden, assisted or unassisted walking), living accommodations (residential or institutionalized), fracture type (intra- or extracapsular), comorbidities (cardiovascular, stroke, respiratory, renal, diabetes), malignant disease, and cognition (Mini-Mental Status Examination). We applied the NHFS, which evaluates seven factors, with scores that range from 0 to 10, as a predictor of 30-day mortality after hip surgery. Results: Survivor scores showed greater variability (CV=0.28) than those of non-survivors (CV=0.20). The receiver operating characteristic curve identified a score of 5.5 as the optimal cutoff point. At this point, the test's sensitivity and specificity indicate the simultaneous maximum likelihood of 30-day survival or non-survival. Conclusion: The NHFS is a robust predictor of 30-day mortality after hip surgery and an updated equation has been validated for the patients in this sample, which increases its clinical credibility.

1. Serviço de Ortopedia e Traumatologia, Hospital Universitário Antônio Pedro. Universidade Federal Fluminense. Niterói, RJ, Brazil.

*Correspondence address:
E-mail: rodalbuquerque19@gmail.com
ORCID: https://orcid.org/0000-0001-8209-1548

BJHBS, Rio de Janeiro, 2024;23(1):36-44
DOI: 10.12957/bjhbs.2024.85386
Received on 20/11/2023. Approved on 29/01/2024.

Keywords: Hip fractures, Risk scoring tools, Mortality, Older adults.

Introduction

Among older adults, hip fractures are one of the most common injuries that require emergency surgery and are among the most common causes of death after an accident.[1]

The Brazilian population is clearly aging: the share of the population over 60 years was 4.2% in in 1950, 8.6% in 2000, 12.1% in 2011, and, according to the Brazilian Institute of Geography and Statistics, was expected to reach 14% in 2020.[2] Aging is a dynamic process that involves an increasing loss of functional reserve and biochemical, morphological, and psychological changes, which makes older adults more prone to environmental risks and, consequently, risk of falls. The mortality rate after hip fracture increases by 4% each year, but when combined with comorbidities it can increase by up to 40%.[3]

The mortality rate in older hip fracture patients has been the subject of considerable discussion,[4-6] with several factors affecting prognosis: age, nutritional status, cognitive status, clinical comorbidities, time between the fracture and surgery, and type of anesthesia. It is unclear whether the most common protocols for identifying patients at higher risk of post-operative complications are suitable for older patients with hip fractures due to their generic and complex measurement systems.[7] The most common preoperative assessment in Brazil is the scale of the American Society of Anesthesiologists, which functions well in the general population, but whose validity is limited in the case of a specific, homogeneous group of patients with hip fracture.[8]

Since 1999, the Queen's Medical Center (Nottingham, UK) has been collecting retrospective data from the charts of its patients with fractures of the proximal femur and is using this information to determine the most important prognostic factors in predicting 30-day mortality risk. This resulted in the Nottingham Hip Fracture Score (NHFS).[7-8]

The NHFS combines previously tested highly significant indices[7-8] that take into account age, sex, hemoglobin and creatinine tests at admission, the presence of cardiovascular, cerebrovascular, respiratory, or renal complications, current or past cancer treatment, living conditions (institutionalization), and a cognitive assessment (Mini-Mental Status Examination) as predictive factors for 30-day postoperative mortality.

The lack of indices to identify high-risk patients among older Brazilians in our health services led us to apply the NHFS initially to a population of hip fracture patients over 65 years of age, correlating the results with their clinical data. However, the aim of the present study was to apply the NHFS prospectively to a cohort of hip fracture patients over 60 years of age to assess whether the NHFS can be used to predict 30-day postoperative mortality.

Material and methods

A total of 216 patients with proximal femur fractures who were treated in the emergency department between March 30, 2016 and March 20, 2018 were analyzed in a non-randomized prospective cohort. The inclusion criteria were hip fractures requiring surgery, age over 60 years, and written informed consent. The exclusion criteria were subtrochanteric hip fractures, pathological fractures, metabolic alterations, congenital deformities, and fracture sequelae. All procedures were conducted per the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study was approved by the Institutional Research Ethics Committee (number 68456417.3.0000.5243).

At admission, patients responded to a questionnaire that assessed age, sex, mobility status (bedridden or assisted or unassisted walking), living accommodations (residential or institutionalized), fracture type (intra- or extracapsular), comorbidities (cardiovascular, cerebrovascular, respiratory, renal, diabetes), and previous malignant disease, after which they completed a cognitive assessment test (Mini-Mental Status Examination).[8-10]

Laboratory blood tests were taken on the admission date, as part of a preoperative routine that included: complete blood count, coagulogram, glucose, urea, creatinine, sodium, potassium, principal components analysis, urine culture, total protein dosage, albumin, globulin, 25-hydroxy-vitamin D, serum calcium, parathyroid hormone, total cholesterol and transferrin.

After data collection and laboratory tests, the NHFS was applied, which assesses 7 factors, with scores ranging from 0 to 10. Patients scoring ⩾ 6 are considered to be high-risk (Table 1).

Table 1. Nottingham Hip Fracture Score

Variables	Values	Points
Age	65-85 years	3
	> 85 years	4
Sex	Male	0
	Female	1
Admission hemoglobin level	< 10 g/dl^{-1}	1
MMSE score	< 6 of 10	1
Institutionalized	Yes	1
Number of comorbidities	> 2	1
Malignancy	Yes	1

Legend: MMSE: Mini-mental Status Examination.
Source: The authors (2023).

To characterize the patient profile, fracture type, and treatment, a descriptive statistical analysis of the behavior of all variables was performed. Data from quantitative variables were summarized using statistical calculations (mean, median, minimum, maximum, standard deviation, coefficient of variation (CV)), simple frequency distributions, and cross tables. A quantitative variable's distribution was considered low if $CV<0.20$; moderate if $0.20 \leq CV<0.40$ and high if $CV \geq 0.40$.

The distribution of patient class frequencies was obtained by determining the number of classes with Sturges' formula, given by $n_c = 1 + 3.32 \log n$, and the range of classes, given by $h = \frac{R}{n_c}$, where R is the total range of the data.

The incidence of mortality was estimated using inferential analysis of the distributions of qualitative variables. The significance of the association between two variables, or the difference between the distributions of proportions, was investigated using the chi-square test. In circumstances where the chi-square results were inconclusive and when possible, Fisher's exact test was used. When a significant association was identified between a factor and 30-day mortality, the estimator used to express the risk was the odds ratio, which evaluated the relationship between the probability that an individual in a group would die in the first 30 days after surgery compared to the probability that an individual in a complementary group would also do so. The significance of the odds ratio was assessed through a confidence interval at the 95% confidence level, which cannot reach 1, since this result would indicate that individuals from both groups have an equal probability of dying in the first 30 days after surgery.

In an inferential analysis of quantitative variables, the independent groups (death and survival) were compared using the Mann-Whitney test due to the small size of the non-surviving group.

When a significant association was found between a quantitative variable and a factor, an optimal cut-off point for it was found using the receiver operating characteristic (ROC) curve methodology. The performance measure for the proposed diagnostic test and cut-off point was the area under the ROC curve (AUC), and the significance of the AUC was evaluated by a test that assesses the null hypothesis H0: AUC=0.5. In addition to this significance test, an asymptotic confidence interval was obtained for the AUC at the 95% confidence level.

In addition to the ROC curve analysis, a logistic regression analysis was performed to validate the NHFS as a predictor of 30-day mortality after surgery. This procedure generated a mathematical model that can predict, based on patient scores, the probability of 30-day mortality.

All analyses were performed considering a maximum significance level of 5% (0.05), i.e., the null hypothesis was rejected whenever the p-value associated with the test was <0.05.

Results

This study is based on a sample of 216 patients with fractures of the proximal femur, treated at a reference orthopedic hospital. Among these patients, 15 died within 30 days after surgery. Table 2 shows the distribution of patient characteristics, both overall and in groups, according to the 30-day survival rate. The characteristics and treatments of the fractures, both overall and in the groups divided by the 30-day survival rate, demonstrate that 94 patients (43.5%) had unstable fractures and 137 (63.4%) had stable fractures. Additionally, 54 patients (25.0%) were treated with dynamic hip screws, 87 (40.3%) with locking rods, 58 (26.9%) by hemiarthroplasty, 12 (5.6%) by total hip arthroplasty, and five (2.3%) with cannulated screws. In total, 113 patients (52.2%) were able to walk without assistance. The majority of patients did not require blood transfusion or blood products (60.6%) and did not experience postoperative complications (78.7%).

Table 2. Distributions according to sex, living accommodations, and age, both overall and in groups according to 30-day survival

Variable		Global n=216		Death in first 30 days				Subgroup comparison p-value
				No n=201		Yes n=15		
Sex	Female	173	80.1%	161	80.1%	12	80.0%	1.000*
	Male	43	19.9%	40	19.9%	3	20.0%	
Living accommodations	Rest home	9	4.2%	9	4.5%	0	0.0%	1.000*
	Residence	207	95.8%	192	95.5%	15	100.0%	
Age	61 to 66	5	2.3%	5	2.5%	0	0.0%	0.001**
	66 to 71	14	6.5%	14	7.0%	0	0.0%	
	71 to 76	18	8.3%	18	9.0%	0	0.0%	
	76 to 81	33	15.3%	33	16.4%	0	0.0%	
	81 to 86	59	27.3%	56	27.9%	3	20.0%	
	86 to 91	50	23.1%	44	21.9%	6	40.0%	
	91 to 96	24	11.1%	19	9.5%	5	33.3%	
	96 to 101	2	0.9%	2	1.0%	0	0.0%	

Legend: * Fisher's exact test ** Mann-Whitney test.

Source: The authors (2023).

The variability in the score was greater in the survivor group than in the non-survivor group, with coefficients of variation of 0.28 and 0.20, respectively. The distribution of scores in the two groups is presented in Figure 1, in which the NHFS distributions were compared using a non-parametric approach, the Mann-Whitney test, due to the small size of the non-survivor group. This test resulted in a p value <0.001, which indicates a significant difference between the score distributions, with higher scores in the non-survivor group. To find an NHFS cutoff point that would serve as an indicator of 30-day postoperative mortality, an analysis of the ROC curve was performed. A score of 5.5 was identified as the ideal cutoff point, at which point sensitivity and specificity reached maximum simultaneous likelihood. This means that, for patients with scores ⩾5.5 (or ⩾6, considering that the NHFS is always a whole number), the death test is capable of predicting mortality within 30 days after surgery.

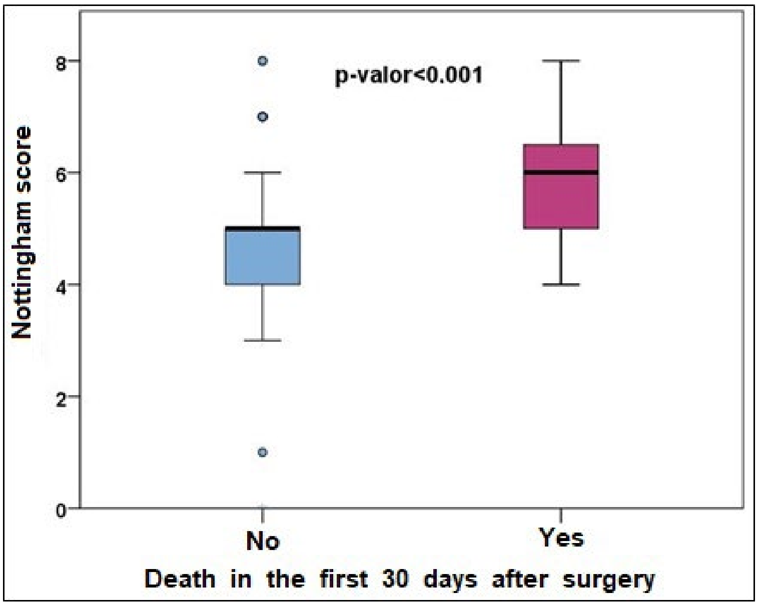

Figure 1. Nottingham Hip Fracture Score distributions according 30-day mortality
Source: The authors (2023).

Table 3 presents the measures of false positives (1-specificity), sensitivity and specificity for different cutoff points, indicating tests with more or less rigorous criteria. Less restrictive cutoff points present greater sensitivity and specificity 1 (points in the upper right corner of the curve), but these cutoff points may result in lower specificity. The ideal cutoff point found was 5.5, which presented an adequate balance between sensitivity and specificity.

For the test based on the cutoff point of 5.5, sensitivity was 0.60, specificity 1 was 0.244, and specificity was 0.756. This means that the test correctly identified 60.0% of patients who died within 30 days but would also have a false positive probability of 24.4%.

A logistic regression model can be used to estimate 30-day mortality based on the NHFS score. The dependent variable of the model is the occurrence of death within 30 days after surgery, while the NHFS score, which showed a significant association with death, is considered to be an independent variable.

Table 3. Nottingham Hip Fracture Score cut-off points for tests with more and less strict criteria

	Cut-off point	1-Specificity	Sensitivity	Specificity
Less strict criteria	-1.00	1.000	1.000	0.000
	0.50	0.990	1.000	0.010
	2.00	0.985	1.000	0.015
	3.50	0.816	1.000	0.184
	4.50	0.547	0.867	0.453
Optimal point	5.50	0.244	0.600	0.756
Stricter criteria	6.50	0.065	0.200	0.935
	7.50	0.010	0.067	0.990
	9.00	00.000	0.000	1.000

Source: The authors (2023).

Figure 2 represents the probability of death after a femoral fracture. Based on this figure, it is possible to predict patient mortality on admission by using the NHFS score. For scores >6, the probability of death is higher than the 30-day mortality rate (6.9%), which was corroborated by the ROC curve analysis.

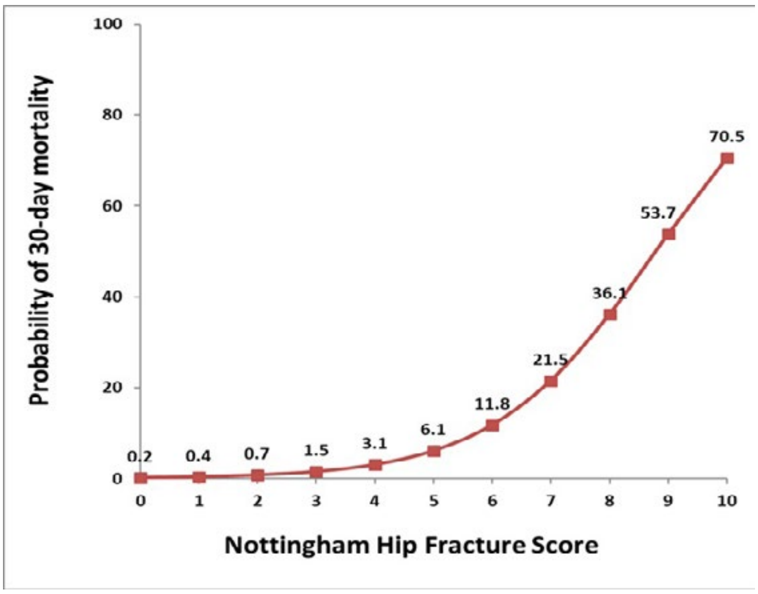

Figure 2. Graphic representation of the probability of death after femur fracture according to Nottingham Hip Fracture Score
Source: The authors (2023).

Discussion

Our study assessed 30-day mortality after hip surgery in 216 hip fracture patients, including patient profiles and treatment, identification of factors associated with 30-day mortality, and validation of the NHFS as a predictor of 30-day mortality. The results indicate that the NHFS is a robust predictor of 30-day mortality.

The NHFS score, developed in 2008 as a predictor of 30-day mortality after hip fracture, shows significantly higher 30-day mortality in high-risk patients (NHFS>4).[7] The equation used to validate the score has been recalibrated twice (2012 and 2015)[5-9] to accommodate the gradual reduction in 30-day mortality after hip fracture. Studies have validated the NHFS' ability to predict 30-day mortality after hip surgery in the UK and in international cohorts.[11-16] Doherty and colleagues[1] concluded that the NHFS is a robust tool for assessing 30-day mortality after hip fracture treatment, demonstrating that the NHFS predicts morbidity well and mortality moderately, but is less effective at predicting length of hospital stay or postoperative complications. A little recalibration was needed to reflect local death rates.

With the objective of determining a simple, economical instrument that is easy to calculate, objective, and precise, Marufu and colleagues[5] conducted a qualitative systematic review of 29 studies that evaluated 25 risk stratification tests. They concluded that the NHFS may be the most appropriate instrument currently available for hip fracture patients. However, more studies are needed to confirm it as the tool of choice for predicting 30-day mortality after hip surgery. Rushton and colleagues[7] suggested that patients with NHFS scores ⩾6 should be considered "high risk", but this must be validated by other studies to identify mortality risk in the small percentage of patients with the highest NHFS scores.

The present prospective study demonstrated that the 30-day mortality rate after hip surgery was 6.9% with a confidence interval of (3.6%, 10.4%). In a sample of 4,967 patients from the United Kingdom, Maxwell and colleagues[8] found a 30-day mortality rate after hip surgery of 7.9%. In a sample of 6,202 patients from the United Kingdom, Wiles and colleagues[4] found a mortality rate of 8.3%. In a multicenter sample of 7,290 patients from the United Kingdom, Moppett and colleagues[11] found a mortality rate of 6.6%. In a sample of 1,079 patients, Rushton and colleagues[7] found a mortality rate of 7.3%. The 30-day mortality rates these authors found after hip surgery in retrospective studies in the United Kingdom did not differ significantly from the rate found in the present study, since all these rates are within our estimated confidence interval.

In our study, 30-day mortality after hip surgery was significantly associated with age, total length of hospital stay, time in ICU, length of hospital stay after surgery, and NHFS score. As a validation study of the NHFS as a predictor of 30-day mortality through ROC curve analysis, our study proposes that a NHFS score of 5.5 is the best cut-off point. That is, if the score is 6 or more, death is predicted within 30 days after surgery. This cut-off point is 1 point higher than that of Maxwell and colleagues[8] who created and validated the NHFS for UK patients. In their study, the AUC was 0.719, slightly lower than that of the present study (0.737). They found a sensitivity of 44.2% and a specificity of 80.8%, although ours were 60.0% and 75.6%, respectively. Thus, adapting the cut-off point to 6 yielded greater diagnostic power for our patients than the cut-off point used in the United Kingdom. Our cut-off point had lower specificity for this population but higher sensitivity in predicting those who would actually die. Since the objective of our study was to correctly predict death, a more sensitive test was considered better than a more specific test, including a lower chance of false negatives. The test of Maxwell and colleagues, despite its higher specificity, had a 55.7% probability of false negatives.[8]

The logistic regression analysis also showed that our model predicted mortality better than that of Maxwell and colleagues,[8] with probabilities >50% for scores of 9 and 10. In Maxwell and colleagues,[8] the logistic model was more conservative and indicated a probability of mortality >50% only for NHFS scores of 10. Comparing our results with those of Maxwell and colleagues,[8] including the ROC curve and the logistic regression model, we conclude that the NHFS is validated for predicting 30-day mortality after hip surgeries in this population. It should be noted that to calculate the NHFS, it was necessary to adapt the version of the Mini-Mental Status Examination commonly used in Brazil,[10] reducing it to a scale of 0 to 10.

This study has some limitations. Although the cohort was prospective, which is important for obtaining accurate data and complete information from patients treated at the institution, the patient sample was small compared to those described in the literature. Also, the study included a single population and was conducted at a single hospital, despite the procedures being performed by three senior surgeons. Limited information may affect the results, and additional studies involving more hospitals could strengthen the evidence. Another limitation was the inclusion of treatment for all proximal femur fractures without differentiation between fractures treated by arthroplasty (total or partial) or osteosynthesis (locking nails, dynamic hip screws, and cannulated screws). This heterogeneity of treatment methods could create a bias and thus affect 30-day postoperative mortality and NHFS results. Despite this, our results are in line with those in the literature (30-day mortality rate, 6.9%).[4,7,8,11] New studies assessing a specific treatment group using the same methods may contribute to stronger and more consistent NHFS results.

Conclusion

The NHFS is a robust predictor of mortality in the first 30 days after hip surgery and, after using an updated equation, was validated for the patients in this sample, increasing its clinical credibility.

References

1. Doherty WJ, Stubbs TA, Chaplin A, et al. Prediction of Postoperative Outcomes Following Hip Fracture Surgery: Independent Validation and Recalibration of the Nottingham Hip Fracture Score. J Am Med Dir Assoc. 2021 Mar;22(3):663-669.
2. Instituto Brasileiro de Geografia e Estatística (IBGE). Agência IBGE. Notícias. Número de idosos cresce 18% em 5 anos e ultrapassa 30 milhões em 2017. Avaiable in: http://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/20980. Accessed 10 January 2018.
3. Ekman EF. The role of the orthopaedic surgeon in minimizing mortality and morbidity associated with fragility fractures. J Am Acad Orthop Surg. 2010 May;18(5):278-85.
4. Wiles MD, Moran CG, Sahota O, et al. Nottingham Hip Fracture Score as a predictor of one year mortality in patients undergoing surgical repair of fractured neck of femur. Br J Anaesth. 2011 Apr;106(4):501-4.
5. Marufu TC, Mannings A, Moppett IK. Risk scoring models for predicting peri-operative morbidity and mortality in people with fragility hip fractures: Qualitative systematic review. Injury. 2015 Dec;46(12):2325-34.
6. Monnerat VB, Ramos AMP, Mathias MB, et al. Avaliação da mortalidade no pós-operatório da fraturas de fêmur em idosos com comorbidades prévias. Fisioter Bras. 2021:22(1):49-60.
7. 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(1):100-3.
8. 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(4):511-7.
9. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-98.
10. Ritter SRF, Zoccoli TLV, Lins MMP, et al. Adaptation of a delirium screening test for elderly adults admitted to emergency departments. Geriatr Gerontol Aging. 2018;12:81-88.
11. Moppett IK, Parker M, Griffiths R, et al. Nottingham Hip Fracture Score: longitudinal and multi-assessment. Br J Anaesth. 2012;109(4):546-50.

12. Karres J, Heesakkers NA, Ultee JM, et al. Predicting 30-day mortality following hip fracture surgery: evaluation of six risk prediction models. Injury. 2015;46(2):371-7.

13. Kau CY, Kwek EB. Can preoperative scoring systems be applied to Asian hip fracture populations? Validation of the Nottingham Hip Fracture Score (NHFS) and identification of preoperative risk factors in hip fractures. Ann Acad Med Singap. 2014;43(9):448-53.

14. Tilkeridis K, Ververidis A, Kiziridis G, et al. Validity of Nottingham Hip Fracture Score in Different Health Systems and a New Modified Version Validated to the Greek Population. Med Sci Monit. 2018 Oct 27;24:7665-7672.

15. De Jong L, Mal Klem T, Kuijper TM, et al. Validation of the Nottingham Hip Fracture Score (NHFS) to predict 30-day mortality in patients with an intracapsular hip fracture. Orthop Traumatol Surg Res. 2019;105(3):485-489.

16. Olsen F, Lundborg F, Kristiansson J, et al. Validation of the Nottingham Hip Fracture Score (NHFS) for the prediction of 30-day mortality in a Swedish cohort of hip fractures. Acta Anaesthesiol Scand. 2021;65(10):1413-1420.