1   **Prediction of postoperative outcomes following hip fracture surgery:**

2   **independent validation and recalibration of the Nottingham Hip Fracture**

3   **Score**

4   William J Doherty*[a], Thomas A Stubbs*[a], Andrew Chaplin[b] MBBS FRCP, Mike R Reed[b] MBBS

5   MD FRCS, Avan A Sayer[c] MBBS PhD, Miles D Witham^[c] BMBCh PhD, Antony K Sorial^[b]

6   MBChB, PGCert, MRes, MRCSEd, MRCSEng

7

8   [a] Campus for Ageing and Vitality, Newcastle University, Newcastle Upon Tyne, NE4 6BE, UK.

9   [b]Department of Trauma and Orthopaedics, Northumbria Healthcare NHS Foundation Trust, Newcastle upon

10   Tyne, NE27 0QJ

11   [c]AGE Research Group, NIHR Newcastle Biomedical Research Centre, Translational and Clinical Research Institute,

12   Faculty of Medical Sciences, Newcastle University and Newcastle Upon Tyne Hospitals NHS Foundation Trust,

13   Newcastle-upon-Tyne, NE4 5PL, UK.

14   *Thomas A Stubbs and William J Doherty contributed equally to this manuscript.

15   ^Miles D Witham and Antony K Sorial are joint senior authors.

16

17   *Correspondence to*:

18   Mr Antony K Sorial, Department of Trauma and Orthopaedics, Northumbria Healthcare NHS

19   Foundation Trust, Newcastle upon Tyne, NE27 0QJ.

20   Email: (Tony.Sorial@newcastle.ac.uk)

21   Professor Miles Witham, AGE Research Group, NIHR Newcastle Biomedical Research Centre,

22   Campus for Ageing and Vitality, Newcastle upon Tyne NE4 5PL.

23   Tel: +44 191 208 1317. Email: Miles.Witham@newcastle.ac.uk

24

25

26   *Running title*:

27   Independent validation of the Nottingham Hip Fracture Score

28

29   *Key words:* Hip fracture; Casemix; Validation; Discrimination; Risk score; Calibration

30

34

35   *Word count:*

36   Abstract 299 words, main text 3000 words, 27 references and 660 words in main

37   tables/figures.

38   Graphics count: 3 tables, 2 figures, 3 supplementary figures, 1 supplementary table.

39

40   *Brief summary:*

41   Independent validation and recalibration of the Nottingham Hip Fracture Score

42   demonstrating moderate ability to predict mortality and functional recovery but poor

43   prediction of postoperative complications.

44

49

50   **Abstract**

51   *Objectives*

52   Independent validation of risk scores after hip fracture is uncommon, particularly for

53   evaluation of outcomes other than death. We aimed to assess the Nottingham Hip Fracture

54   Score (NHFS) for prediction of mortality, physical function, length of stay and postoperative

55   complications.

56

57   *Design*

58   Analysis of routinely collected prospective data.

59

60   *Setting and Participants*

61   Consecutive hip fracture patients were identified from the Northumbria hip fracture database

62   between 2014-2018. Patients were excluded if they were not surgically managed or if scores

63   for predictive variables were missing.

64

65   *Methods*

66   C-statistics were calculated to test the discriminant ability of the NHFS, Abbreviated Mental

67   Test Score (AMTS), and ASA grade for in-hospital, 30- and 120-day mortality, functional

68   independence at discharge, 30-days and 120-days, length of stay, and postoperative

69   complications.

70

71   *Results*

72   We analysed data from 3208 individuals, mean age 82.6 (SD 8.6). 2192 (70.9%) were female.

73   194 (6.3%) died during the first 30-days, 1686 (54.5%) were discharged to their own home,

74   211 (6.8%) had no functional mobility at 120-days, 141 (4.6%) experienced a postoperative

75   complication. The median length of stay was 18 days (IQR 8-28). For mortality, c-statistics for

76   the NHFS ranged from 0.68-0.69, similar to ASA and AMTS. For functional independence, the

77   c-statistics for the NHFS ranged from 0.74-0.83, similar to AMTS (0.61-0.82) and better than

78   the ASA grade (0.68-0.71). Length of stay was significantly correlated with each score

79   (p<0.001 by Jonckheere-Terpstra test); NHFS and AMTS showed inverted U-shaped

80   relationships with length of stay. The best predictor of postoperative complications was ASA

81   grade (0.56-0.64), whilst NHFS (0.49-0.59) and AMTS (0.48-0.57) performed similarly poorly.

82

83   *Conclusions and Implications*

84   The NHFS performed consistently well in predicting functional outcomes, moderately in

85   predicting mortality, but less well in predicting length of stay and complications. There

86   remains room for improvement by adding further predictors such as measures of physical

87   performance in future analyses.

88

89   **Introduction**

90   Hip fractures are the most common serious injury in older people, the most common reason

91   for older people to need emergency surgery, and the most common cause of death following

92   an accident. In 2017, there were over 65000 hip fractures in the UK.[1] More people are

93   surviving to old age, with accompanying frailty and multi-morbidity; frailty is a syndrome of

94   impaired homeostasis and vulnerability to minor stressors. 14% of those aged 60 and over

95   are living with frailty in England.[2] In England, hip fracture prevalence peaked for males in the

96   early-2000s and for females in the mid-1990s;[3] however, evolving demographics means that

97   patients have increasingly complex medical, social and rehabilitation needs.[4] The ability to

98   predict mortality after hip fracture surgery is vital for informing prognostic conversations and

99   adjusting for casemix when benchmarking service performance. The ability to predict other

100   outcomes such as functional mobility, length of stay, and postoperative complications are

101   also important to patients and healthcare services.

102   The Nottingham Hip Fracture Score (NHFS) was developed in 2008 to predict mortality at 30-

103   days following hip fracture.[5] The equation used to derive variable weights has been

104   recalibrated twice, in 2012 and 2015, to accommodate a gradual reduction in overall 30-day

105   mortality after hip fracture.[6,7] The NHFS uses a simple scoring system, which includes age,

106   gender, haemoglobin on admission, malignancy, comorbidities, Abbreviated Mental Test

107   Score (AMTS) and preoperative residential status. All the variables are objective measures

108   and are routinely checked or known at admission, reducing observer and omission bias, and

109   facilitating easy calculation. It is scored out of 10, with a score of 10 being associated with

110   the highest risk of mortality.

111   The ability of the NHFS to predict 30-day mortality has been validated independently and

112   internationally.[8-10] There have been several validations of the NHFS in predicting mortality in

113   the UK, with some revealing limitations when applied to differing demographics or subsets of

114   the cohort.[6,7] Far less research has focussed on the NHFS in predicting other important

115   outcomes, including length of stay and complications. Few studies have sought to test the

116   NHFS in predicting functional outcomes[11] and one-year mortality,[12] and there is no published

117   independent validation in predicting these outcomes. The ability of the NHFS to predict

118   postoperative complications has been neglected by the literature.

119   As the NHFS was purpose-built to predict 30-day mortality, it remains unclear whether the

120   tool reliably predicts a broader range of outcomes of importance to patients, surgical teams

121   and healthcare organisations. Therefore, this study aims to provide independent validation

122   of the NHFS with regards to mortality, and assess its predictive ability for functional

123   outcomes, length of stay and postoperative complications.

124

125   **Methods**

126   *Analysis population*

127   Data were prospectively gathered on consecutive hip fracture patients attending

128   Northumbria Healthcare NHS Foundation Trust between 1 October 2009 and 1 January 2019.

129   All in-patient data were directly entered into a local database by specialist nurses and

130   subsequently uploaded to the National Hip Fracture Database (NHFD).[13] NHFS and AMTS

131   were calculated at admission by emergency medicine and orthopaedic speciality doctors.

132   American Society of Anaesthesiologists (ASA) grade was evaluated by the anaesthetic team

133   in the immediate pre-operative period. The NHFS was calculated for patients from 1 April

134   2014 onwards. Post-discharge outcomes were assessed through follow-up phone interviews

135   conducted by specialised NHFD nurses directly with patients, relatives or institutions to which

136   the patients were discharged. Data on medical complications were obtained from routine

137   clinical records collected prospectively on all in-patients at the trust. The study was managed

138   in accordance with Caldicott principles. As the analysis did not require any new patient

139   contact or data collection, the study did not require evaluation by an ethics committee.

140   Datasets were linked using the patient's trust ID number and anonymised within the trust

141   before export for analysis. Individuals were excluded from analysis if their data could not be

142   linked, if data on NHFS, AMTS or ASA grade were incomplete or missing, or if the patient was

143   managed non-operatively.

144   The AMTS comprises 10 questions and is used to assess cognitive impairment and composes

145   part of the NHFS. An AMTS of ≥8 is deemed normal. We assessed the ability of the AMTS

146   alone to predict outcomes, as a simple alternative to the NHFS. The AMTS is widely used,

147   simple to collect, and requires no additional medical history or biochemical results, making it

148   suitable for use in resource-poor settings. The AMTS has been demonstrated as an

149   independent predictor of length of hospital stay following hip fracture[14] and eventual return-

150   to-home.[11] The AMTS can identify patients suffering from delirium,[15] which is a frequent

151   complication in hip fracture associated with longer hospital stays, increased rates of other

152   complications and mortality.[16]

153   We also compared the NHFS to ASA grade, which scores patients from I (healthy) to VI (brain-

154   dead). It is a subjective global measure associated with length of stay following hip fracture

155   surgery, severity of complications and total hospitalisation costs.[14,17,18]

156

157   *Selection of postoperative outcomes*

158   The NHFS was designed to predict 30-day mortality.[5] Other outcomes are also important,

159   including functional mobility, postoperative complications, and length of stay; however, there

160   is no consensus on which tool to use to predict such outcomes. The outcomes we assessed

161   included mortality, physical function, length of stay and complications.

162   We assessed all-cause mortality at three intervals: within 30-days post-surgery, within 120-

163   days post-surgery and in-hospital. These variables were selected to provide a clear

164   comparison against the existing NHFS literature, whilst allowing for calibration of the NHFS

165   for our dataset. For physical function we analysed residential status after discharge and self-

166   reported mobility. Residential status was analysed by the number of patients at home or

167   sheltered housing within 30-days and within 120-days post-surgery: an assessment of social

168   independence following discharge. Mobility was analysed by the number of patients

169   reportedly freely mobile outdoors without aids at 30-days and 120-days post-surgery, mobile

170   outdoors with or without aids at 30-days and 120-days post-surgery, or with no functional

171   mobility at 30-days and 120-days post-surgery. 'No mobility' was defined as being wheelchair

172   or bed-bound. Length of stay was derived from hospital admission and discharge data; where

173   patients were transferred to rehabilitation wards/units after surgery, the date of discharge

174   from the rehabilitation unit was used instead. To assess the rate of postoperative

175   complications, we used the following records from the database: deep vein thrombosis (DVT)

176   or pulmonary embolism (PE) within 60-days post-surgery, urinary tract infection (UTI) or

177   pneumonia within 30-days post-surgery, stroke or myocardial infarction (MI) within 30-days

178   post-surgery and any post-operative complication (excluding DVT and PE) within 30-days

179   post-surgery. Postoperative complications and mobility variables were selected due to their

180   availability within the dataset and their utility in assessing a broad range of outcomes.

181

182

183

Statistical Analysis

Data were analysed using SPSS v25 (IBM, New York, USA). A two-sided p value of <0.05 was taken as significant for all analyses. Descriptive data were generated for baseline characteristics of the study dataset.

Analysis of discriminant ability of predictive variables

Receiver-operator characteristic (ROC)-curves and c-statistics with 95% confidence intervals were generated to compare the performance of the NHFS, AMTS and ASA grade in predicting mortality, post-discharge residence, mobility and postoperative complications. The relationship between predictor variables and length of stay was assessed using the Jonckheere-Terpstra test. Boxplots were generated to demonstrate the relationship between length of stay and the predictor variables.

Recalibration

To recalibrate the original NHFS equation to best fit the observed 30-day mortality for our cohort, binary logistic analysis was performed on the same patients involved in the discriminant analysis. A binary variable was produced for 30-day mortality following hip fracture surgery (deceased=0; living=1), which was used as the dependent variable and the NHFS as the independent variables. The values generated by the regression analysis were inserted into an equation in the format of "Risk (%) = $\frac{100}{1 + e^{C+X(NHFS)}}$", where "C" is a constant and "X" is a coefficient of the NHFS. This equation is modelled on the format of the original equation produced in 2008[5] and its subsequent revisions.[6,7] A 95% confidence interval was calculated using odds ratios produced by the regression analysis.

207   **Results**

208   NHFS values were available for 3208 patients. Due to missing NHFS (n=30), AMTS values (n=8)

209   and non-operative management (n=78), further patients were excluded (*see Figure 1*).

210   Baseline details (n=3092) are shown in Table 1, and the prevalence of each outcome is shown

211   in Table 2. Supplementary Table 1 shows the prevalence of each outcome by NHFS score,

212   demonstrating the expected gradients of mortality, postoperative function and

213   complications.

214

215   *Mortality*

216   At 30-days post-surgery, 194 (6.3%) patients had died, with a mean age of 86.1 years (SD 8.0);

217   107 (55.2%) were women. At 120-days (n=1503) post-surgery, 278 (18.5%) were dead. There

218   were 243 (7.9%) deaths in-hospital, with a mean age of 86.0 years (SD 8.1); 142 were women

219   (58.4%). The NHFS had moderate discriminative ability for mortality at 30-days, 120-days and

220   in-hospital, with c-statistics of 0.690 (0.652-0.727), 0.676 (0.641-0.710) and 0.684 (0.650-

221   0.718) respectively. ASA grade had similar discriminative ability for 30-day and in-hospital

222   mortality, but performed better in predicting 120-day mortality, with a c-statistic of 0.702

223   (0.669-0.735). AMTS had the poorest discriminative ability for mortality, with c-statistics

224   ranging from 0.614 (0.579-0.648) to 0.627 (0.591-0.662), shown in Table 3.

225

226   *Residential status*

227   At 30-days post-surgery, 1330 (43.1%) had returned to their own home or sheltered housing.

228   At 120-days post-surgery, 777/1503 (51.7%) had returned to their own home or sheltered

229   housing. The NHFS had good discriminative ability for return-to-home / sheltered housing

230   both at 30- and 120-days, with c-statistics of 0.769 (0.752-0.785) and 0.793 (0.770-0.816)

231   respectively); ASA grade was poorer, but AMTS outperformed the NHFS at both 30-days

232   (0.793 (0.778-0.809)) and 120-days (0.823 (0.802-0.845)), shown in Table 3.

233

234   *Postoperative mobility*

235   At 120-days post-surgery, 211 (6.8%) patients had no functional mobility. NHFS had moderate

236   discriminative ability for functional outcomes, with c-statistics ranging from 0.672 (0.617-

237   0.728) (in predicting 'no functional mobility at 120-days') to 0.834 (0.722-0.946) (in predicting

238   'freely mobile outdoors without aids at 30-days'). Similar results were seen for both ASA grade

239   and AMTS, with c-statistics between 0.654 (0.595-0.713) to 0.876 (0.807-0.945) and 0.609

240   (0.443-0.774) to 0.788 (0.745-0.831) respectively, shown in Table 3.

241

242   *Length of Stay*

243   Significant differences in length of stay at different scores was discovered for NHFS, AMTS

244   and ASA grade, but both NHFS and AMTS showed inverted U-shaped relationships between

245   score and length of stay. The relationship between the predictor variables and median length

246   of stay is demonstrated in Figure 2.

247

248   *Postoperative complications*

249   For DVT/PE, c-statistics were comparable for NHFS (0.536 (0.452-0.620)) and ASA grade

250   (0.529 (0.457-0.602)), which were superior to AMTS (0.502 (0.422-0.583)). For UTI, NHFS

251   (0.589 (0.550-0.628)) and AMTS (0.577 (0.538-0.616)) were comparable and were superior to

252   ASA grade (0.536 (0.498-0.574)). For pneumonia, ASA grade (0.614 (0.543-0.684)), NHFS

253   (0.585 (0.510-0.661)) and AMTS (0.562 (0.496-0.627)) were comparable. NHFS (0.571 (0.498-

254   0.644)) scored similarly to ASA grade (0.614 (0.540-0.688)) and AMTS (0.519 (0.452-0.586)) in

255   predicting Stroke/MI. An aggregate variable for all complications (excluding DVT and PE)

256   within 30-days was created, with c-statistics demonstrated in Table 3.

257   *Recalibration*

258   Binary logistic regression was used to recalibrate the NHFS equation for 30-day mortality. The

259   recalibrated equation used a constant of -5.020 and a coefficient of 0.434. This formed the

260   equation "Risk (%) = $\dfrac{100}{1 + e^{(0.434*NHFS)-5.02}}$ ". The relationship between observed and predicted

261   values for the 2015 NHFS and the recalibrated equation, and a comparison between the

262   predicted mortality rates for each point of the NHFS, are shown in Supplementary Figures 1,

263   2 and 3.

264

265   **Discussion**

266   Using prospectively gathered data, we have independently validated the NHFS as a predictor

267   of postoperative outcomes in hip fracture surgery. The NHFS performed consistently well in

268   discriminating in-hospital mortality, and mortality, residential status and mobility at 30- and

269   120-days post-surgery. A significant association was evident between NHFS and length of

270   stay, but the score performed poorly in predicting postoperative complications. The NHFS

271   showed good agreement between observed and predicted 30-day mortality in our sample,

272   with little change on recalibration. We observed that length of stay was lowest at both

273   extremes of the NHFS, a phenomenon also seen with the AMTS. This is likely because patients

274   with the highest NHFS (and lowest AMTS) are more often admitted from residential care;

275   consequently, a lower degree of functional recovery is required before discharge, expediting

276   the process compared to those with intermediate scores.

277   The ability of the NHFS to predict 30-day mortality has been the focus of most prior studies.

278   This has been validated in both UK and international cohorts, where 30-day mortality is

279   significantly increased in high-risk patients (NHFS >4).[9,12] Only one study has examined the

280   NHFS in predicting return-to-home, finding that increasing NHFS negatively correlates with

281   eventual return-to-home.[11] This is also the only study to evaluate the NHFS in predicting

282   length of stay, showing that increasing NHFS negatively correlates with the proportion of

283   patients returning to home at 7-, 14- and 21-days.[11] No previous studies have examined the

284   ability of the NHFS to predict independence of mobility and postoperative complications.

285   The original cohort of patients in which the NHFS was designed had a mean age 2.8 years

286   (3.4%) younger, 5.5% fewer males and 1.6% higher 30-day mortality than this cohort.[5] Whilst

287   increasing age and male sex are risk factors for postoperative mortality,[19,20] the improved

288   survival rate observed in our population may be attributed to the introduction of the BPT

289   (Best Practice Tariff) in 2010.[21] The BPT uses monetary incentives to ensure standards are met

290   and provide the best quality of care. BPT criteria include time to surgery (<36 hours) and

291   assessment by a geriatrician within 72 hours. Mortality has been consistently decreasing since

292   the adoption of the BPT, with 30-day mortality decreasing from 6.9% to 6.1% from 2017 to

293   2018.[13] These changes are likely to contribute to improved survival rate observed in our

294   sample compared to that predicted by previous calibrations.

295   The NHFS performs consistently well in predicting mortality and functional outcomes, but

296   poorly in predicting postoperative complications. Clinical benefit may be derived from

297   broadening the predictive scope of the NHFS to include post-operative complications. This

298   may be achieved through combining the NHFS with another validated outcome predictor;

299   though AMTS and ASA are practical options, as they are routinely collected, neither had

300   superior performance to the NHFS across the outcome measures. Alternatively, a novel

301   variable could be collected; a potential candidate is handgrip strength, which may add

302   predictive value through measuring domains of pathophysiology that are not captured within

303   the NHFS, namely sarcopenia. Handgrip strength is easy to collect, strongly aligned to function

304   across multiple body systems and predicts all-cause mortality in older people.[22] Unlike tests

305   of lower limb function, handgrip strength can be measured preoperatively for hip fracture

306   patients, even in the presence of cognitive impairment. Initial data suggests that handgrip

307   strength predicts functional outcomes after hip fracture,[23,24] but no studies have assessed

308   whether it can improve the discriminant abilities the NHFS when used in combination.

309   Another alternative is to incorporate a frailty measurement; these have been operationalised

310   in several different ways, including via a phenotypic score including grip strength,[25] a

311   cumulative deficit score that can be collected electronically,[26] and a simple pictorial

312   assessment of preoperative function.[27]

313

314   *Implications for clinical practice and research*

315   Our results confirm the utility of the NHFS in predicting postoperative mortality,[6,7] and extend

316   the external validation of the NHFS as a predictor of functional outcomes.[7] The NHFS is

317   therefore useful to adjust for casemix in these outcomes at an organisation level, and to

318   inform prognostic discussions with patients. The NHFS has modest utility in adjusting for

319   casemix in length of stay statistics and is not well suited in adjusting for casemix in terms of

320   postoperative complications. Future research could focus on comparing other easy to collect

321   tools, such as handgrip strength or frailty measures, with the NHFS in predicting a range of

322   outcomes, or adding these to the NHFS to improve its predictive ability. Having well-

323   calibrated outcome prediction tools for hip fracture surgery are important in resource

324   management; in an economically strained healthcare system, with one of the lowest ratios of

325   hospital beds/population in the western world, the ability to predict in-hospital mortality,

326   discharge destination and length of stay allow for bed management and aid clinical decisions

327   made by orthopaedic and geriatric medicine teams.

328

329   *Strengths and limitations*

330   Strengths of the study include use of a large patient cohort with prospectively collected data,

331   and examination of multiple important outcomes rather than mortality alone. Although this

332   study included 3092 patients, some were excluded due to missing data on postoperative

333   complications (n=116), and only half of patients had data at 120-days. 120-day follow up was

334   conducted by telephone, thus responses directly from patients could have been limited by

335   cognitive impairment, ill health, change in place of permanent residence, and death. We did,

336   however, collect data from patient relatives, and in the case of patients in institutions, from

337   trained staff and carers, which could ameliorate a proportion of this effect. There was a large

338   discrepancy in ASA grade results between our dataset and the 2015 validation dataset;[6] fewer

339   than 1% of patients in the 2015 dataset were ASA grade 1 or 2 vs 29.5% in our dataset. The

340   large discrepancy in ASA grade highlights the inconsistent, subjective nature of the

341   assessment; reports of high numbers of patients with ASA grade 1 or 2 is a systemic issue

342   seen at units across the UK.[13] The single-centre nature of this study limits the generalisability

343   of results; although the study area contains varied levels of socioeconomic deprivation, the

344   older population in the catchment of Northumbria is overwhelmingly white, limiting

345   generalisability to other ethnic groups.

346

347

348

349   **Conclusion and Implications**

350   The Nottingham Hip Fracture Score performs well at predicting mortality and functional

351   outcomes but is less effective at predicting length of stay or postoperative complications. A

352   modest recalibration was required to reflect local mortality rates, but our results confirm that

353   the score is useful in predicting mortality and in predicting who will return to independent

354   living. Additional measures may improve the ability of the NHFS to predict other important

355   outcomes – particularly length of stay and postoperative complications. Simple measures that

356   reflect multisystem body function or frailty, for instance handgrip strength, could provide this

357   additional information and merits further investigation.

358

359   *Conflicts of interest*

360   The authors declare no competing interests.

361

362

363

364

365

366

367

368

369

370

371

372

373   *Tables and figures*

374   **Table 1. Preoperative characteristics of 3092 included patients.**

Variable	
Age, mean (SD) (years)	82.7 (8.6)
Female sex (%)	2192 (70.9%)
ASA grade	
1	76 (2.5)
2	838 (27.1)
3	1752 (56.7)
4	426 (13.8)
NHFS	
0 (%)	60 (1.9)
1	57 (1.8)
2	33 (1.01)
3	374 (12.1)
4	736 (23.8)
5	769 (24.9)
6	583 (18.9)
7	358 (11.6)
8	93 (3.0)
9	24 (0.8)
10	5 (0.2)
Median (IQR)	5 (4-6)
Mean (SD)	4.9 (1.6)
AMT score	
Median (IQR)	8 (3-10)
Mean (SD)	6.6 (3.8)
Pre-fracture Residential status (%)	
Nursing care	117 (3.8)
Other hospital site of this trust	40 (1.3)
Other hospital trust	2 (0.1)
Own home/sheltered housing	2338 (75.6)
Residential care	570 (18.4)
This hospital site	25 (0.8)
Pre-fracture Mobility status (%)	
Freely mobile without aids	974 (31.5)

Mobile outdoors with one aid	519 (16.8)
Mobile outdoors with two aids or frame	337 (10.9)
No functional mobility	105 (3.4)
Some indoor mobility but never goes outside without help	1156 (37.4)
Pathological fracture (%)	
Atypical	2 (0.1)
Atypical bisphosphonate type subtrochanteric fracture	5 (0.2)
Malignancy	43 (1.4)
Non-pathological	3041 (98.4)

375

376   NHFS: Nottingham Hip Fracture Score. AMT: Abbreviated Mental Test. ASA: American Society of

377   Anaesthesiologists.

378

379   **Table 2. Prevalence of post-surgical outcomes in the Northumbria hip fracture cohort (n=3092).**

Post-surgical outcomes	N (%)
Mortality	
Mortality at 30-days	194/3083 (6.3)
Mortality at 120-days	278/1503 (18.5)
In-hospital mortality	243/3087 (7.9)
Discharge destination	
Own home or sheltered housing at 30-days	1330/3083 (43.1)
Own home or sheltered housing at 120-days	777/1503 (51.7)
Postoperative mobility	
Freely mobile outdoors without aids at 30-days	6/829 (0.7)
Freely mobile outdoors without aids at 120-days	93/881 (10.6)
Mobile outdoors with/without aids at 30-days	122/829 (14.7)
Mobile outdoors with/without aids at 120-days	254/881 (28.8)
No functional mobility at 30-days	319/829 (38.5)
No functional mobility at 120-days	81/881 (9.2)
Postoperative complications	
Deep venous thrombosis or pulmonary embolism within 60-days	47/1908 (2.5)
Urinary tract infection within 30-days	215/1908 (11.3)
Pneumonia within 30-days	55/1908 (2.9)
Stroke/myocardial infarction within 30-days	47/1908 (2.5)

Any postoperative complication (excluding DVT/PE) within 30-days	726/1908 (38.1)

380

381   DVT: Deep venous thrombosis. PE: Pulmonary embolism

382   **Table 3. Discriminant statistics for outcomes predicted by NHFS, AMTS and ASA grade.**

Variables	n	c-statistics [95% CI]		
		NHFS	AMTS	ASA grade
Mortality				
Mortality at 30-days	3083	.690 [.652-.727]	.624 [.585-.663]	.695 [.658-.732]
Mortality at 120-days	1503	.676 [.641-.710]	.627 [.591-.662]	.702 [.669-.735]
In-hospital Mortality	3087	.684 [.650-.718]	.614 [.579-.648]	.699 [.667-.732]
Discharge destination				
Own home/sheltered housing at 30-days	3083	.769 [.752-.785]	.793 [.778-.809]	.688 [.669-.707]
Own home/sheltered housing at 120-days	1503	.793 [.770-.816]	.823 [.802-.845]	.713 [.688-.739]
Postoperative mobility				
Freely mobile outdoors without aids at 30-days	829	.834 [.722-.946]	.609 [.443-.774]	.876 [.807-.945]
Freely mobile outdoors without aids at 120-days	881	.766 [.718-.814]	.676 [.621-.731]	.739 [.683-.795]
Mobile outdoors with/without aids at 30-days	829	.779 [.739-.820]	.785 [.749-.820]	.741 [.691-.790]
Mobile outdoors with/without aids at 120-days	881	.737 [.704-.770]	.750 [.717-.783]	.678 [.641-.715]
No functional mobility at 30-days	829	.693 [.656-.729]	.707 [.670-.743]	.653 [.616-.690]
No functional mobility at 120-days	881	.672 [.617-.728]	.788 [.745-.831]	.654 [.595-.713]

| Postoperative complications				
DVT/ PE within 60-days	46	.536 [.452-.620]	.502 [.422-.583]	.529 [.457-.602]
Urinary tract infection within 30-days	214	.589 [.550-.628]	.577 [.538-.616]	.536 [.498-.574]
Pneumonia within 30-days	54	.585 [.510-.661]	.562 [.496-.627]	.614 [.543-.684]
Stroke/ myocardial infarction within 30-days	47	.571 [.498-.644]	.519 [.452-.586]	.614 [.540-.688]
Any postoperative complication (excluding DVT and PE) within 30-days	342	.579 [.547-.611]	.561 [.529-.592]	.568 [.537-.599]

383

384   NHFS: Nottingham Hip Fracture Score. AMTS: Abbreviated Mental Test Score. ASA: American Society
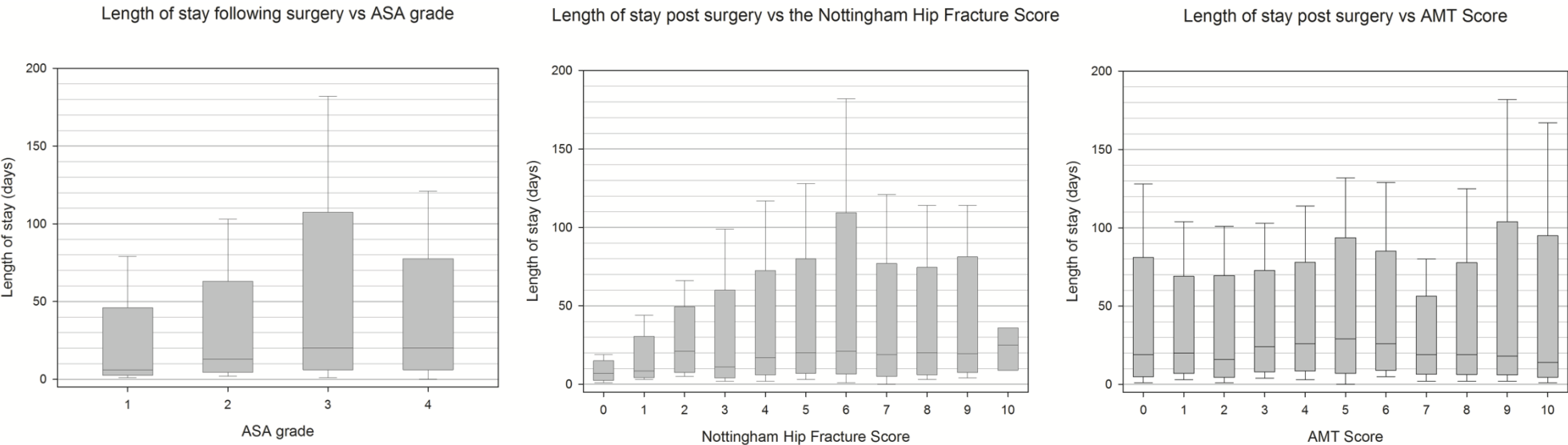
385   of Anaesthesiologists. DVT: Deep venous thrombosis. PE: Pulmonary embolism

386

387   **Figure 1. Flowchart of included patients.**
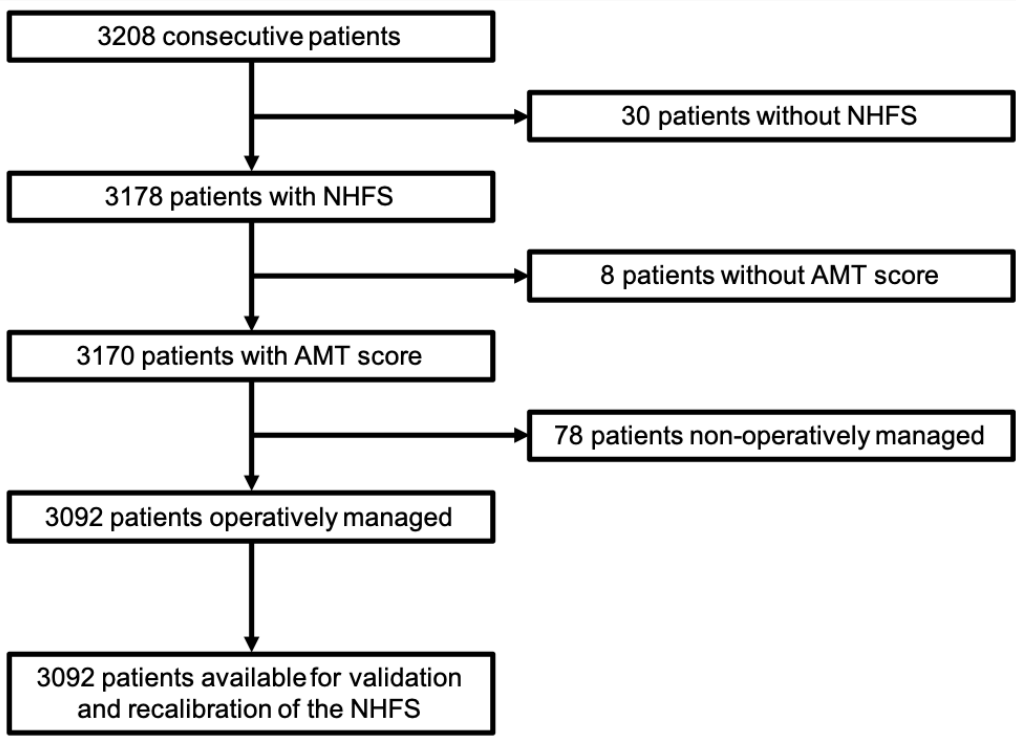
390 **Figure 2. Relationship between predictor variables and median length of stage (n=1908).**

391

392   p<0.001 by Jonckheere-Terpstra test   p<0.001 by Jonckheere-Terpstra test   p<0.001 by Jonckheere-Terpstra test

393

394   **References**

395   1.   Royal College of Physicians. National Hip Fracture Database annual report 2018.

396   https://www.nhfd.co.uk/files/2018ReportFiles/NHFD-2018-Annual-Report-v101.pdf. Accessed on

397   16th April, 2020.

398   2.   Gale C, Cooper C, Sayer A. Prevalence of frailty and disability: findings from the English

399   Longitudinal Study of Ageing. Age and Ageing. 2014;44(1):162-5.

400   3.   Lucas R, Martins A, Severo M, et al. Quantitative modelling of hip fracture trends in 14

401   European countries: testing variations of a shared reversal over time. Scientific Reports.

402   2017;7(1):3754.

403   4.   Baker P, Salar O, Ollivere B, et al. Evolution of the hip fracture population: time to consider

404   the future? A retrospective observational analysis. BMJ Open. 2014;4.

405   5.   Maxwell MJ, Moran CG, Moppett IK. Development and validation of a preoperative scoring

406   system to predict 30 day mortality in patients undergoing hip fracture surgery. Br J Anaesth.

407   2008;101(4):511-7.

408   6.   Marufu TC, White SM, Griffiths R, et al. Prediction of 30-day mortality after hip fracture

409   surgery by the Nottingham Hip Fracture Score and the Surgical Outcome Risk Tool. Anaesthesia.

410   2016;71(5):515-21.

411   7.   Moppett IK, Parker M, Griffiths R, et al. Nottingham Hip Fracture Score: longitudinal and multi-

412   centre assessment. British Journal of Anaesthesia. 2012;109(4):546-50.

413   8.   Karres J, Heesakkers NA, Ultee JM, Vrouenraets BC. Predicting 30-day mortality following hip

414   fracture surgery: Evaluation of six risk prediction models. Injury. 2015;46(2):371-7.

415   9.   Kau CY, Kwek EB. Can preoperative scoring systems be applied to Asian hip fracture

416   populations? Validation of the Nottingham Hip Fracture Score (NHFS) and identification of

417   preoperative risk factors in hip fractures. Ann Acad Med Singapore. 2014;43(9):448-53.

418   10.   Tilkeridis K, Ververidis A, Kiziridis G, et al. Validity of Nottingham Hip Fracture Score in

419   Different Health Systems and a New Modified Version Validated to the Greek Population. Medical

420   science monitor : international medical journal of experimental and clinical research. 2018;24:7665-

421   72.

422   11.   Moppett IK, Wiles MD, Moran CG, Sahota O. The Nottingham Hip Fracture Score as a predictor

423   of early discharge following fractured neck of femur. Age and Ageing. 2011;41(3):322-6.

424   12.   Wiles MD, Moran CG, Sahota O, Moppett IK. Nottingham Hip Fracture Score as a predictor of

425   one year mortality in patients undergoing surgical repair of fractured neck of femur. Br J Anaesth.

426   2011;106(4):501-4.

427   13.   Royal College of Physicians. National Hip Fracture Database annual report 2019

428   https://www.nhfd.co.uk/files/2019ReportFiles/NHFD_2019_Annual_Report_v101.pdf. Accessed on

429   27th April, 2020.

430   14.   Richards T, Glendenning A, Benson D, et al. The independent patient factors that affect length

431   of stay following hip fractures. Ann R Coll Surg Engl. 2018;100(7):556-62.

432   15.   Pendlebury ST, Klaus SP, Mather M, et al. Routine cognitive screening in older patients

433   admitted to acute medicine: abbreviated mental test score (AMTS) and subjective memory complaint

434   versus Montreal Cognitive Assessment and IQCODE. Age and Ageing. 2015;44(6):1000-5.

435   16.   Mosk CA, Mus M, Vroemen JP, et al. Dementia and delirium, the outcomes in elderly hip

436   fracture patients. Clin Interv Aging. 2017;12:421-30.

437   17.   Kastanis G, Topalidou A, Alpantaki K, et al. Is the ASA Score in Geriatric Hip Fractures a

438   Predictive Factor for Complications and Readmission? Scientifica. 2016;2016:7096245-.

439   18.   Michel JP, Klopfenstein C, Hoffmeyer P, et al. Hip fracture surgery: is the pre-operative

440   American Society of Anesthesiologists (ASA) score a predictor of functional outcome? Aging Clin Exp

441   Res. 2002;14(5):389-94.

442   19.   Endo Y, Aharonoff GB, Zuckerman JD, et al. Gender differences in patients with hip fracture: a

443   greater risk of morbidity and mortality in men. J Orthop Trauma. 2005;19(1):29-35.

444   20.   Wang AQ, Ng BH, Cheung LP, Chin RP. Factors Affecting Mortality and Hospital Admissions

445   after Hip Surgery among Elderly Patients with Hip Fracture in Hong Kong - Review of a Three-Year

446   Follow-Up(). Hong Kong journal of occupational therapy : HKJOT. 2017;30(1):6-13.

447   21.   NHS England, NHS Improvement. 2019/20 National Tariff Payment System - A consultation

448   notice: Annex   DtD   2019.

449   https://improvement.nhs.uk/documents/484/Annex_DtD_Best_practice_tariffs.pdf.   Accessed   on

450   16th April, 2020.

451   22.   García-Hermoso A, Cavero-Redondo I, Ramírez-Vélez R, et al. Muscular Strength as a Predictor

452   of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis

453   of Data From Approximately 2 Million Men and Women. Archives of Physical Medicine and

454   Rehabilitation. 2018;99(10):2100-13.e5.

455   23.   Di Monaco M, Castiglioni C, De Toma E, et al. Handgrip strength is an independent predictor

456   of functional outcome in hip-fracture women: a prospective study with 6-month follow-up. Medicine

457   (Baltimore). 2015;94(6):e542-e.

458   24.   Orwig D, Magaziner J, Fielding RA, et al. Application of SDOC Cut-points for Low Muscle

459   Strength for Recovery of Walking Speed After Hip Fracture. The Journals of Gerontology: Series A.

460   2020. https://doi.org/10.1093/gerona/glaa076

461   25.   Fried LP, Tangen CM, Walston J, et al. Frailty in Older Adults: Evidence for a Phenotype. The

462   Journals of Gerontology: Series A. 2001;56(3):M146-M57.

463   26.   Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using

464   routine primary care electronic health record data. Age and Ageing. 2016;45(3):353-60.

465   27.   Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in

466   elderly people. Canadian Medical Association Journal. 2005;173(5):489.