Prediction of 90-day mortality after Total Hip Arthroplasty: a simplified and externally validated model based on observational registry data from Sweden, England and Wales

Anne Garland\*1,2,3, Erik Bulow\*2,4, Erik Lenguerrand5, Ashley Blom5, Mark Wilkinson6, Adrian Sayers5, Ola Rolfson2,4, Nils P. Hailer1,2

2020-03-24

# Affiliation and contact

1. Department of Orthopaedics, Institute of Surgical Sciences, Uppsala University Hospital, Uppsala, Sweden
2. Swedish Hip Arthroplasty Register, Gothenburg, Sweden
3. Department of Orthopaedics, Visby Hospital, Visby, Sweden
4. Department of Orthopaedics, Institute of Clinical Sciences, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
5. Translational Health Sciences, Bristol Medical School, University of Bristol, UK.
6. Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK.

\* Authors contributed equally.

**Correspondence:** [erik.bulow@vgregion.se](mailto:erik.bulow@vgregion.se). Postal: Svenska Höftprotesregistret, Registercentrum Västra Götaland, SE-413 45, Sweden

##### PAGE BREAK

# Abstract

**OBJECTIVE:** Shared individualized decision making prior to total hip arthroplasty (THA) includes discussions of the risk of mortality. However, diagnosis code-based instruments such as the Charlson co-morbidity index are impossible to use in clinical practice, and the American Society for Anesthesiologists (ASA) class is imprecise. We thus developed a simple model to predict 90-day mortality after elective THA.

**PATIENTS AND METHODS:** We studied 53,099 Swedish patients operated with a cemented THA due to primary osteoarthritis 2008 - 2015 and used personal identity numbers to link data from the Swedish Hip Arthroplasty Register, the National Patient Register, the Longitudinal integrated database for health insurance and the population register. We used a bootstrap ranking procedure together with logistic LASSO regression to develop a prediction model for death within 90 days after surgery. The model was internally validated with bootstraping, and externally validated on a cohort of 125,428 patients from the National Joint Registry for England and Wales. Model performance was assessed by the area under the receiver operation characteristics curve (AUC) for discrimination, and by calibration belt plots.

**RESLUTS:** The unadjusted cumulative 90-day mortality was 0.33 % (95 % CI: 0.28 - 0.38) in Sweden and 0.52 % (95 % CI: 0.48-0.56) in England and Wales. We propose a simple main effects model combining age, sex, ASA class, the presence of cancer, disease of the central nervous system, kidney disease, and diagnosed obesity. This model had good discriminatory ability both internally (AUC = 0.78, 95% CI: 0.75 to 0.81) and externally (AUC = 0.75, 95% CI: 0.73 to 0.76), rendering it superioir to traditional models based on the Charlson (AUC = 0.66, 95% CI: 0.62 to 0.70), and the Elixhauser co-morbidity indices (AUC = 0.64, 95% CI: 0.59 to 0.68). Our model was well calibrated for predicted probabilities up to 5 %. A web calculator to aid clinical usage was published at <https://shpr.registercentrum.se/om-registret-1/forskning/prediktion-av-90-dagarsmortalitet/p/SkyeTsTFB>.

**CONCLUSION:** We found a relatively simple prediction model of 90-day mortality after elective cemented THA based on easily available demographic data and the presence of four defined co-morbidities. This may facilitate individualized risk assessment prior to one of the most common surgical interventions.

##### PAGE BREAK

# Introduction

Shared decision making has evolved into an integral part of patient-physician interactions prior to surgical interventions, and the weighing of risks against benefits is central to this process. In terms of gained quality of life and cost-utility, total hip arthroplasty (THA) is an enormously successful procedure. 90-day mortality after THA performed due to osteoarthritis is low, ranging from 0.2% to 0.6% in large joint registry studies.1,2 The risk of short-term mortality is nevertheless an essential part of a preoperative discussion between patient and surgeon.

Co-morbidity is associated with a shorter remaining life span, but the Charlson and Elixhauser co-morbidity indices poorly predict mortality after THA.2,3,3–6 Additionally, these complex co-morbidity instruments are based on information from extensive in- or out-patient databases, each measure being defined by 1,178 and 1,516 individual codes respectively from the Swedish adaptation of the international classification of diseases version 10 (ICD-10). Each co-morbidity index also exists in numerous versions, making interpretation and comparisons difficult.7,8 The use of these indices is therefore limited to research settings but are impossible to use under clinical circumstances.

Pre-operative co-morbidity data have been used to develop clinically applicable prediction models for early mortality and adverse events such as surgical site infection in the context of THA surgery.9–15 The first prediction model based on a European cohort was the Patient Decision support tool for joint remplacement in the UK (<https://jointcalc.shef.ac.uk/>). Due to limited, often single-center based samples, a lack of reported calibration abilities, the absence of external validation, and the use of extrapolation, there is still room for improved prediction models of early mortality.16

We therefore aimed to develop a parsimonious prediction model of 90-day postoperative mortality after THA with internal and external validation of discrimination and calibration, and to compare this to the accuracy of prediction models based on existing co-morbidity measures.

# Patients and Methods

We used data from the Swedish Hip Arthroplasty Register (SHAR) for model derivation and internal validation of discrimination and calibration. The best model was then validated externally on patients from England and Wales recorded in the National Joint Registry for England, Wales, Northern Ireland, the Isle of Man and the States of Guernsey (NJR). We focused on patients with cemented THA due to osteoarthritis. Cementation is the most common fixation technique used in Sweden, thus used as inclusion criteria to increase heterogeneity. Only the last operated hip was accounted for in patients with bilateral THA, since those are shown to better resemble unilateral THAs.17 (Figure 1).

## Derivation cohort (Sweden)

The patient cohort was identified from the SHAR. The inclusion period started 2008 since this was the first year with the American Society for Anesthesiologists (ASA) class and Body Mass Index (BMI) systematically recorded. The last year was 2015, since we had access to co-morbidity from the national patient register up to this point. Deterministic data linkage was performed by 10-digit identity numbers, assigned to all Swedish residents at birth or immigration.18 We used data from a variety of sources, as previously described by Cnudde et al.19 Age, sex, BMI, ASA class, type of hospital (university/county/rural/private) and year of surgery were collected from the SHAR, with a completeness of 96-98%.20 Data on educational level (low = up to 9 years/middle = 10-12 years/high = at least 12 years) was recorded for more than 98 % of the population with 85 % accuracy in the longitudinal integration database for health insurance and labor market studies (LISA) from Statistics Sweden.21 Civil status (married/un-married/divorced/widow[er]), were also collected from LISA. The Swedish National Patient Register (NPR) was used to assess co-morbidity during the year preceding index surgery. This register contains all relevant diagnoses coded by ICD-10, as well as dates of admission and discharge for in- and outpatient episodes in all private and public hospitals. Completeness for NPR is above 99 % and 85-95 % of all diagnostic codes are valid.22 Death dates were linked from the national population register.

Patients with BMI either missing or above 50 were excluded, as were patients with ASA class missing or above III, as well as patients with unknown educational level or type of hospital.

## External validation cohort (England and Wales)

Similar inclusion criteria were applied to the validation cohort from England and Wales. We did not exclude patients based on BMI, educational level or civil status however, since those variables were not used in the final model. The validation cohort was created by linking NJR to the secondary uses service database from the National Health Service (NHS). This database includes co-morbidity recorded in the admitted patient care module of the hospital episode statistics (HES) database in England, and the patient episode database for Wales.23–27 Out-patient deaths were linked from the Office of national statistics. Data linkage used NHS-numbers, as well as patient names, age, sex and address.28

## Defining co-morbidity

Co-morbidity was defined by individual ICD-10 codes grouped into 17 categories according to Charlson29–31 and 31 categories according to Elixhauser.31,32 Some co-morbidities were identified by both Charlson and Elixhauser, and some distinct co-morbidities were closely related (such as hypertension with and without complications, or abuse of either drugs or alcohol). We combined individual diagnostics groups to establish 21 broader categories of co-morbidity (Table 1). Conditions were merged according to clinical relevance as to be recognized in a patient-physician meeting without access to external register data.

## Statistical analysis for model development

We used the Kaplan-Meier estimator to assess unadjusted mortality after cemented THA as background knowledge. Further analysis was based on logistic regression, since no censoring occurred in the Swedish cohort within the first 90 days, and we assumed the same was true for patients from England and Wales. We used a modelling procedure with bootstrap ranking and a logistic least absolute shrinkage and selection operator (LASSO).33,34 Numeric variables (age and BMI) were normalized before modelling to have mean = 0 and standard deviation = 1. This way, the magnitude (absolute values) of their estimated coefficients would indicate variable importance on the same scale as the categorical variables. Co-morbidities recorded for at least one patient who died within 90 days, and one who did not, were included in the modelling process. 1,000 bootstrap samples were drawn from the observed data set.35 We used 10-fold cross validation for every bootstrap sample with a broad range of potential penalty values (:s) in a logistic LASSO model. We then only kept :s minimizing the mean cross-validated deviances in each sample. Those :s were used to estimate model coefficients for each potential predictor. The magnitude (absolute values) of those estimates were used as a measure of variable importance. Piece-wise linear regression was used to detect a break point where a significant drop in variable importance was observed. Potential predictors with variable importance above this break point were considered important and kept as model candidates. The whole process was repeated 100 times. Covariates that were selected at least once out of the 100 times were used in a main effects model of multivariable logistic regression without penalty, and without pre-normalization of numeric variables (main model). A reduced model with variables chosen at least 33 out of the 100 times was used as a simpler alternative for comparison. This model, without cancer as a predictor, was also evaluated separately, considering that medical indications for THA surgery may look different for patients with cancer compared to cancer free patients. Univariable models with the ASA class, Charlson or Elixhauser co-morbidity indices were used for benchmarking, as well as a multivariable model with age and sex. Each model including age was fitted three times, once with age as a main effect and twice with restricted cubic splines, either by two or three knots. Odds ratios for the final model were estimated with 95% confidence intervals.36

## Statistical analysis for model validation

Each model was used to predict the probability of death within 90 days for patients from the SHAR (internal validation). Sensitivity and specificity were estimated to form receiver operating characteristic (ROC) curves and the area under those curves (AUC) were used as a measure of discriminative ability.37 Models with a lower 95% confidence limit above 0.7, were considered good. Those intervals were based on percentiles from 2,000 non-parametric bootstrap samples. We used the bias-corrected Somers’ rank correlation based on 100 resamples to adjust for optimism.38 Calibration bands were made to graphically assess model calibration, comparing predicted probabilities and observed proportions.39 The reduced model was then evaluated externally. An AUC with 95% confidence interval was calculated for the model as-is. Re-calibration of the model intercept was then performed to account for different mortality rates in Sweden compared to England and Wales. An updated over-all slope was also calculated to account for country-specific treatment differences.40 Calibration for this re-calibrated model was illustrated in the same calibration belt plot as for the internal calibration.

## Statistical tools

We built an online web calculator available at <https://shpr.registercentrum.se/om-registret-1/forskning/prediktion-av-90-dagarsmortalitet/p/SkyeTsTFB> to be used in clinical practice.

We used R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) with significant packages tidyverse, tidymodels, furrr, pROC and shiny. All R-scripts and necessary configurations (but no personal data) is available at zenodo.org/XXX **[Add!]**. A linked Binder environment is also available for interactive usage.

## Ethical approval

Ethical approval for this study was obtained from the Regional Ethical Review Board in Gothenburg (360-13).

# Results

## Study participants

53,099 patients (age 35 - 99, 61% females) were included in the derivation cohort from SHAR (Figure 1 left panel). 175 (0.33%) of those patients died within 90 days and no one was censored before that. The unadjusted risk of 90-day mortality was therefore 0.33 % (95 % CI: 0.28 - 0.38). Further characteristics of the study population are presented in Table 2. 26% and 48% of the Swedish patients had at least one pre-surgery co-morbidity according to Charlson and Elixhauser. The proportion of patients with ASA class III was 18%.

In addition, 125,428 patients were included for the external validation cohort (Figure 1 right panel). Their unadjusted risk of 90 day mortality was higher compared to the derivation cohort 0.52 % (95 % CI: 0.48-0.56).

## Model development and internal validation

There were 5 co-morbidities that were not recorded for any patient who died in the Swedish derivation cohort: acquired immunodeficiency syndrome by the human immunodeficiency virus (AIDS/HIV), coagulopathy, fluid electrolyte disorders, liver disease and weight loss. Those variables were therefore excluded as potential predictors prior to any statistical model derivation. The derived main model included age, sex, ASA class, the presence of cancer, disease of the central nervous system, kidney disease, diagnosed obesity, heart condition, anemia, and heart infarct. The reduced model, with covariates included at least 33 out of 100 times, were restricted to age, sex, ASA class, the presence of cancer, disease of the central nervous system, kidney disease, and diagnosed obesity (Table 3).

There were no differences between models including age as a main effect, or as modeled by restricted cubic splines with either two or three knots. We will therefore focus on the simpler models with age as a main effect. Similarly, the correction for optimism only affected the third decimals of the AUC confidence intervals. We will therefore omit those.

The main and reduced models were no different regarding discriminative power, (AUC = 0.79, 95% CI: 0.75 to 0.82 versus AUC = 0.78, 95% CI: 0.75 to 0.81). We therefore considered the reduced model as superior due to simplicity. Traditional models performed poorly with 95 % confidence intervals not above 0.7: the Charlson co-morbitiy index (AUC = 0.66, 95% CI: 0.62 to 0.70), and the Elixhauser com-morbidity index (AUC = 0.64, 95% CI: 0.59 to 0.68) (Figure 2 left panel and 3).

The ability of the reduced model to estimate probabilities of death within 90 days is further illustrated in Figure 4. Patients who died had, on average, higher predicted probabilities to do so. The model calibration was good for estimated probabilities up to 3% and acceptable up to 5%, although with predicted probabilities usually higher then observed proportions of deaths among patients with corresponding covariate patterns (Figure 5).

Estimated coefficients and corresponding odds ratios for the reduced model are presented in Table 4.

To omit cancer from the reduced model did not affect the AUC or calibration for estimated probabilities below 3%. It did lead to worse calibration outside this range however, wherefore we recommend including cancer as a predictor.

## External validation

The discriminative ability for the reduced model was not statistically significantly different when applied to the external validation cohort (AUC = 0.75, 95% CI: 0.73 to 0.76) compared to the internal (AUC = 0.78, 95% CI: 0.75 to 0.81) (Figure 2 right panel and Figure 3). Calibration of the re-calibrated model was slightly worse compared to the internal calibration. Predicted probabilities between 0.5 % and 1.5 % were lower than observed and estimated 95 % confidence bands were wider. Over-all, calibration was still good however for predicted probabilities below 3% and acceptable below 5% (Figure 5).

# Discussion

We found that a multivariable main effects logistic regression model with age, sex, ASA class, the presence of cancer, disease of the central nervous system, kidney disease, and diagnosed obesity better discriminated patients who died within 90 days after THA, from patients who survived, compared to traditional models with the Charlson or Elixhauser co-morbidity indices.

The resulting (reduced) model predicts the probability of death within 90 days as: where is the vector of estimated coefficients presented in table 4. This formula is considered valid for patients aged 35 - 99 years and for predicted probabilities up to 5%.

## Practical usage

Our model could be used in clinical practice, either by the formula above, or by a simple web calculator online (<https://shpr.registercentrum.se/om-registret-1/forskning/prediktion-av-90-dagarsmortalitet/p/SkyeTsTFB>). For example a 35 year old woman with ASA class I and none of the important co-morbidities would have a 0.0028 % risk to die within 90 days of surgery. Another woman, 67 years old (the first age quantile), would have an elevated risk of 0.037 %. A 99 years old man (the maximum observed age) with ASA class III and cancer, would have a risk of 8.8 %. Note however that covariate patters with observed probabilities above 5% were rare (0.13%, n = 70). Estimated mortality risks above 5% are therefore subject to extrapolation. The observed proportions of deaths for patients with similar characteristics is likely lower (as indicated by Figure 5). Some risk calculators ignore this problem,9,15 but we think this should be acknowledged.

## Model predictors

Variables in our model were chosen based on statistical properties and not due to clinical relevance. Variables and estimated coefficients should therefore not be assigned any exact epidemiological and/or causal meaning per se.41 We could nevertheless discuss some aspects of the relevant predictors.

Age and sex are well-known predictors of remaining life span. It is less obvious that this relation must be linear. We used restricted cubic splines to allow a more flexible relation, but found no difference in AUC compared to simpler main effect models. ASA class III has the largest estimated coefficient among all predictors, indicating large relative importance. This is clinically reasonable since a label of “severe systemic disease” should be based on a relevant patient assessment prior to surgery. Conditions were not too severe however since elective surgery was performed after this assessment. ASA class is nevertheless known to have a high degree of internal variability.42 It has previously been compared to the Charlson co-morbidity index, but not with respect to mortality after THA.43,44 Patients with ASA class IV-VI were excluded since those categories describe severe disease, moribund and brain-dead individuals. We suspect that patients with such conditions might have been misclassified. Their true class could either be lower, or they could have received THA due to a femoral hip fracture instead of osteoarthritis.

Obesity was not statistically significant by itself ( 0.05) but was still relevant as a predictor due to unobserved heterogeneity.45 We noted that the proportion of patients with BMI above 30, the WHO definition of obesity is much higher (23.23%) than the proportion of patients with ICD-10 = E66 (diagnosed obesity) from the National Patient Register (1.88%). This proportion corresponds to patients with BMI of at least 39. Thus, the relevance of obesity might not be the condition itself, but rather all circumstances making the related diagnose appear in the national patient register. This might include unmeasured confounders related to patients, doctors, and local hospital guidelines or similar.

## Strengths and limitations

A strength of the study is the nationwide design with data from national registers in both Sweden, England and Wales. The Swedish registers are valid with low proportions of missing data.20,22,46,47 Some concerns have been raised regarding validity of the hospital episodes statistics (HES) database from England and Wales however.48–50 A systematic review found that the overall median diagnostic accuracy (comparing ICD-codes from HES to individual case notes) was 80%.24 Both in- and outpatient co-morbidity data was available for Sweden, but only in-patient data for England and Wales. We did not have acces to any data from primary care or nursing homes in any of the counties.

The linkage procedure between NJR and HES was previously described by Smith et al.51 They found that privately founded patients were not included in HES. 17% of the patients had missing personal data or did not allow linkage, and 6% were not found in HES even if their data was available from NJR.

In addition to data validity, the indication for different fixation techniques also differ in Sweden compared to England and Wales. Cementation is most commonly used in Sweden, but not so in England and Wales, where relatively younger patients are more likely operated without cement. To only include cemented THA therefore implies an over-representation of older and frailer patients from England and Wales compared to Sweden. Those patients might be more likely to die within 90 days. This was seen in the calibration plot where estimated probabilities below 1.5% underestimated the observed proportion of deaths in the external validation.

We investigated wether the effect of age on mortality was non-linear using splines. A similar approach might be relevant for BMI as well, which we only modelled as a linear main effect. A U-shaped association between BMI categorized as underweight (below 18.5 kg/m2), normal (18.5 to 25 kg/m2), or obese (above 25 kg/m2) has been showed for example in England and Wales.52 Our recordings of BMI might be imprecise however.53 Some h○ospitals might report actual measurements but it is unknown to what extent those data are self-estmiated by the patients. Therefore, we aimed to avoid overfitting BMI based on too elaborate modelling.

It should be noted that the risk model does not study THA as an observed intervention. We merely followed the cohort who already had THA. Hence, deaths within 90 days might occur for those patients regardless if THA is inserted or not. The proximity in time however, the maximum of 90 days from THA to death, is an indication that the operation might have influenced the observed mortality. Insertion of an elective THA is always preceded by a clinical judgement. Hence, no patient with a foreseen death near-by is given elective THA. We therefore believe that at least a non-significant proportion of deaths within 90 days are related to the THA itself.

## Conclusion

Our results indicate that the risk of early postoperative mortality after cemented THA can be pre-operatively assessed by a parsimonious prediction model. We hope that this model, with its accompanying web calculator, will facilitate a shared decision making between physicians and their patients in need of THA.

# Contribution of authors

**[ADD NJR CONTRIBUTIONS!]** AG and NH initiated the study and managed the ethical review board application. EB developed the statistical model. EL and AS performed external validation with data from NJR. AG and EB drafted the manuscript. All authors edited and finalized the manuscript.

# Acknowledgement

We would like to thank Szilard Nemes, previous senior statistician at the Swedish Hip Arthroplasty Register, for involvement in planning and interpreting the study.

##### PAGE BREAK

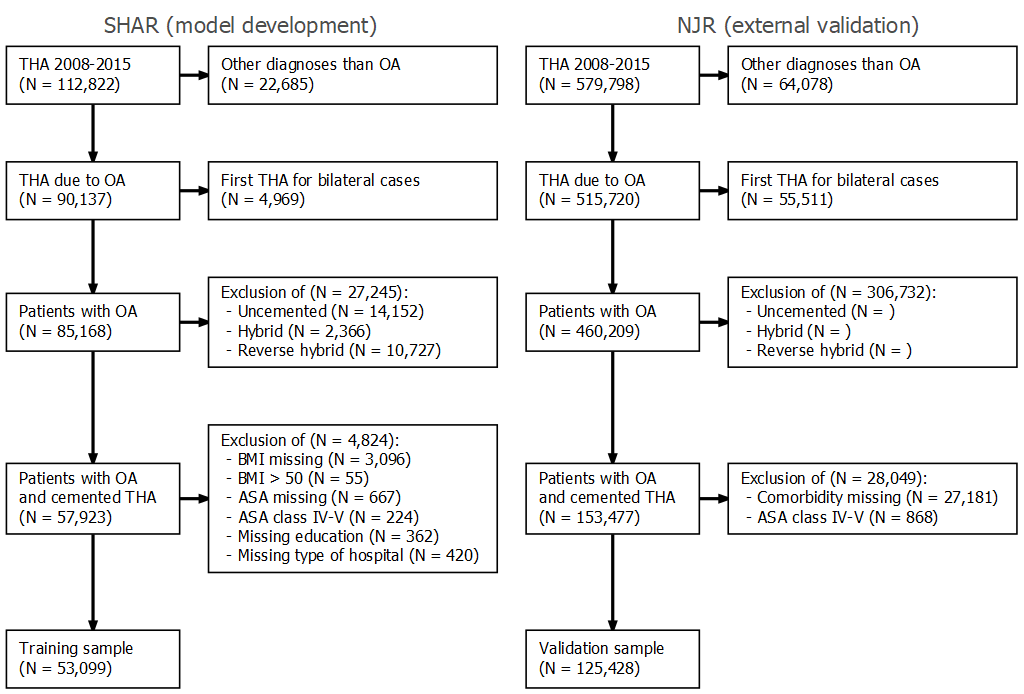


Figure 1: Flowchart depicting inclusion criteria and number of patients. Data from the Swedish Hip Arthroplasty Register was used for model derivation and internal validation (left). Data from the National Joint Registry in England and Wales was used for external validation (right).

##### PAGE BREAK

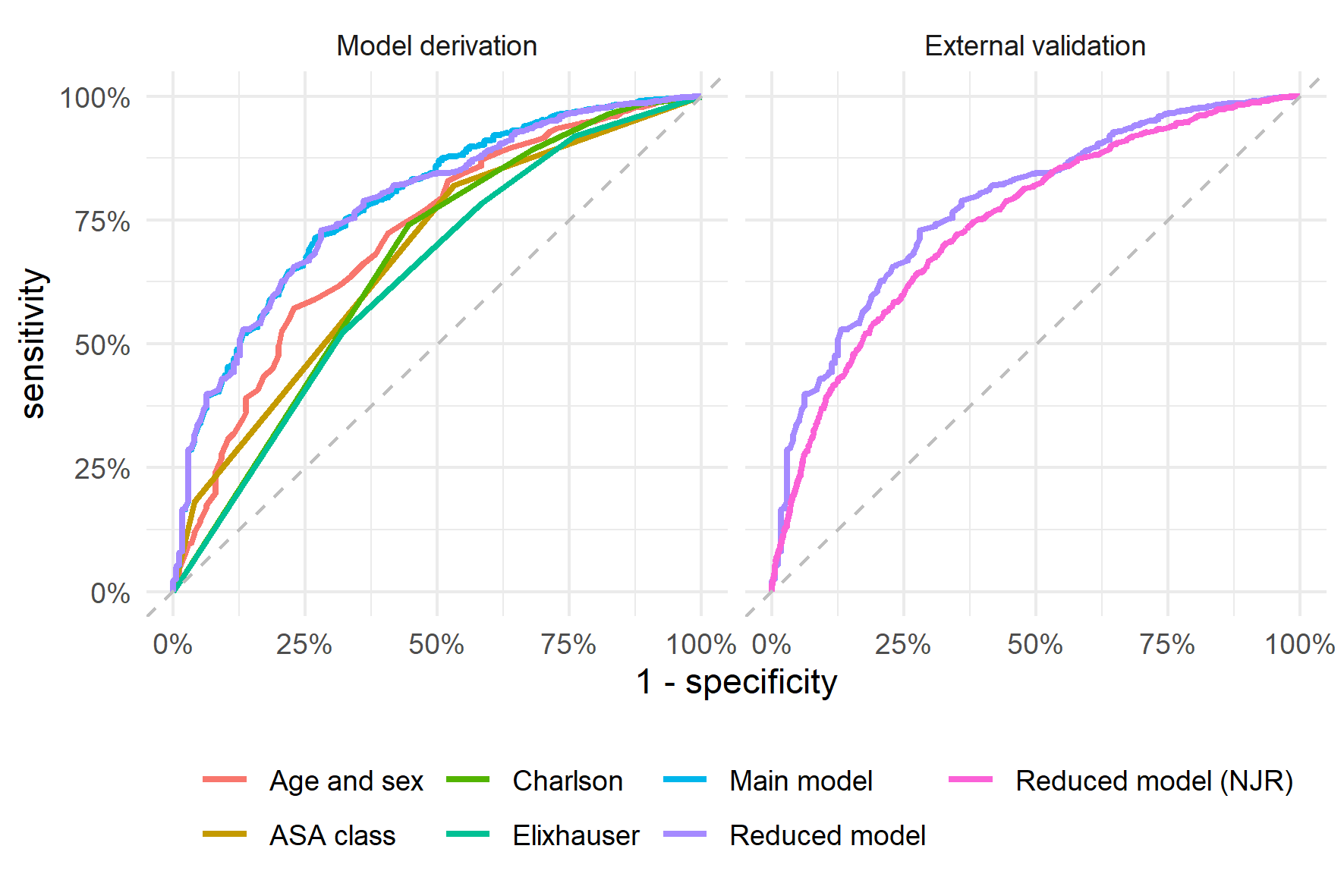


Figure 2: Receiver Operation Characteristics (ROC) curves. The area under the ROC curve (AUC) for the reduced model was no different compared to the main model (left panel). The reduced model works almost as good for predictions in the National Joint Registry for England and Wales (right panel).

##### PAGE BREAK

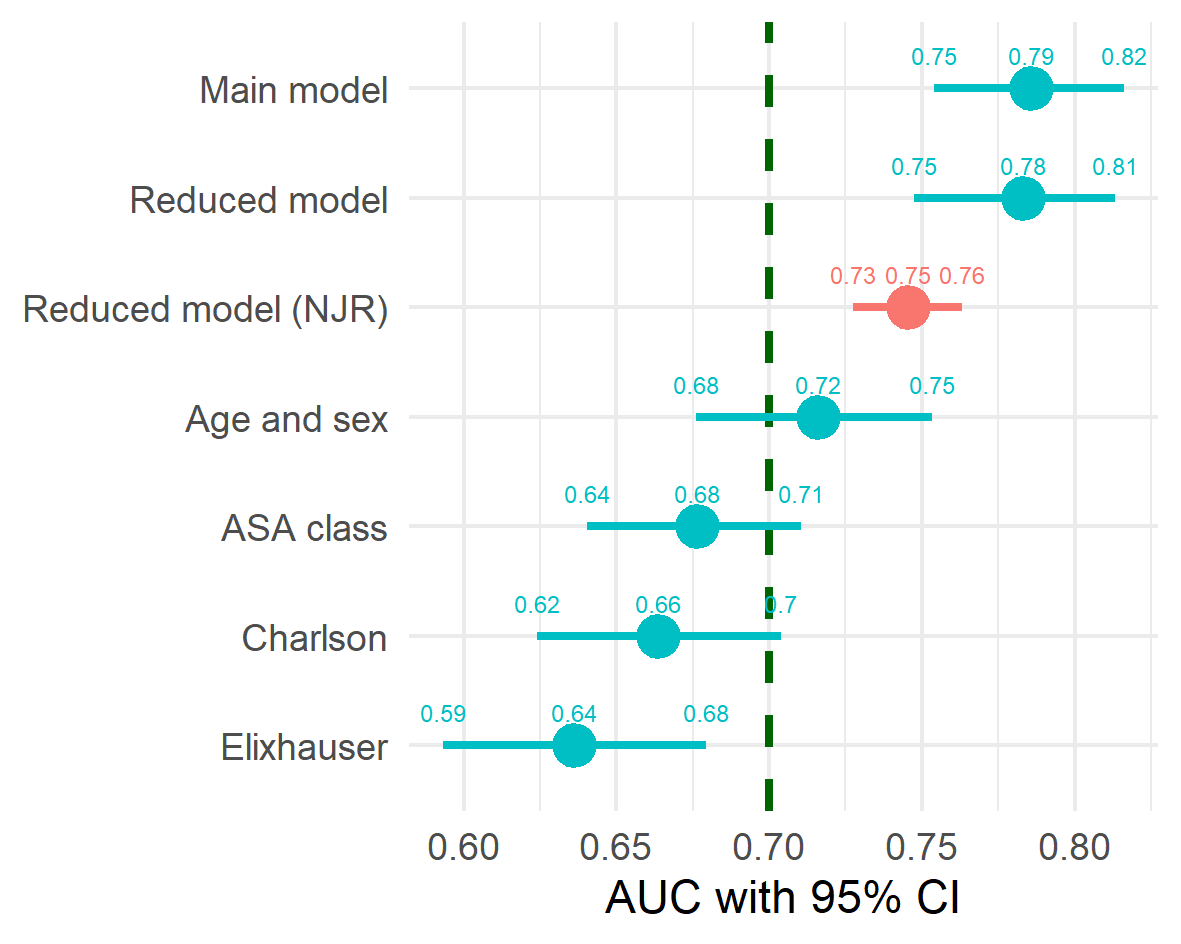


Figure 3: Area Under the Receiver Operation Characteristics Curve (AUC) as a measure of predictive discriminative ability with 95 % bootstrap confidence intervals. The reduced model was no different compared to the main model, neither internally, nor externally (overlapping confidence intervals). AUC above 0.7 was considered good. This was observed for the derived models but not for any other model included for comparison.

##### PAGE BREAK

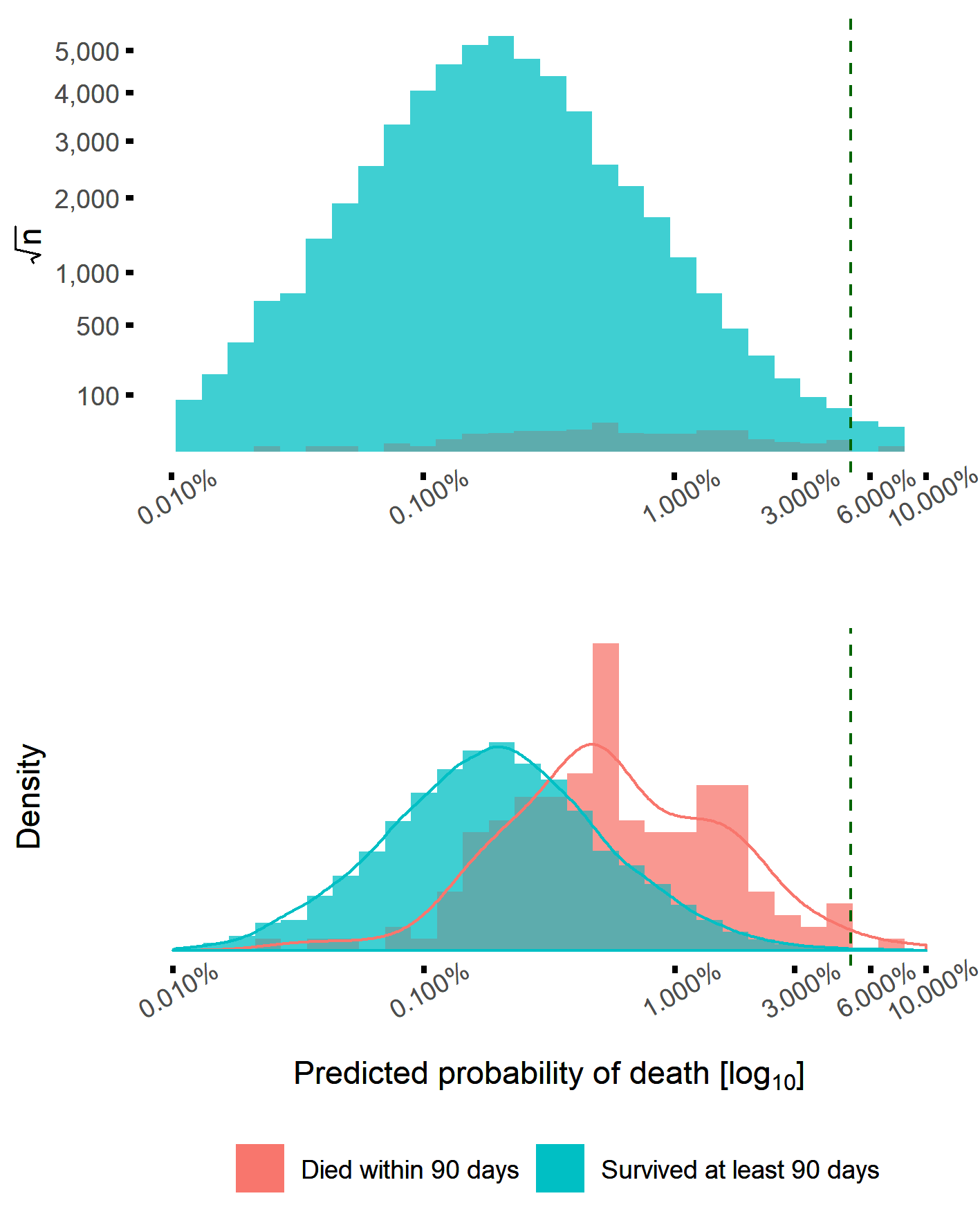


Figure 4: Most patients survived more than 90 days after THA, wherefore the blue bars totally dominate the histogram (upper panel; note the scales). A normalized density plot however reveals that patients who died within 90 days were, on average, estimated to have a higher probability to do so (lower panel). There were very few observations with covariate patterns resulting in death probabilities higher than 5% (green line). Estimated probabilities above this limit are therefore subject to extrapolation.

##### PAGE BREAK

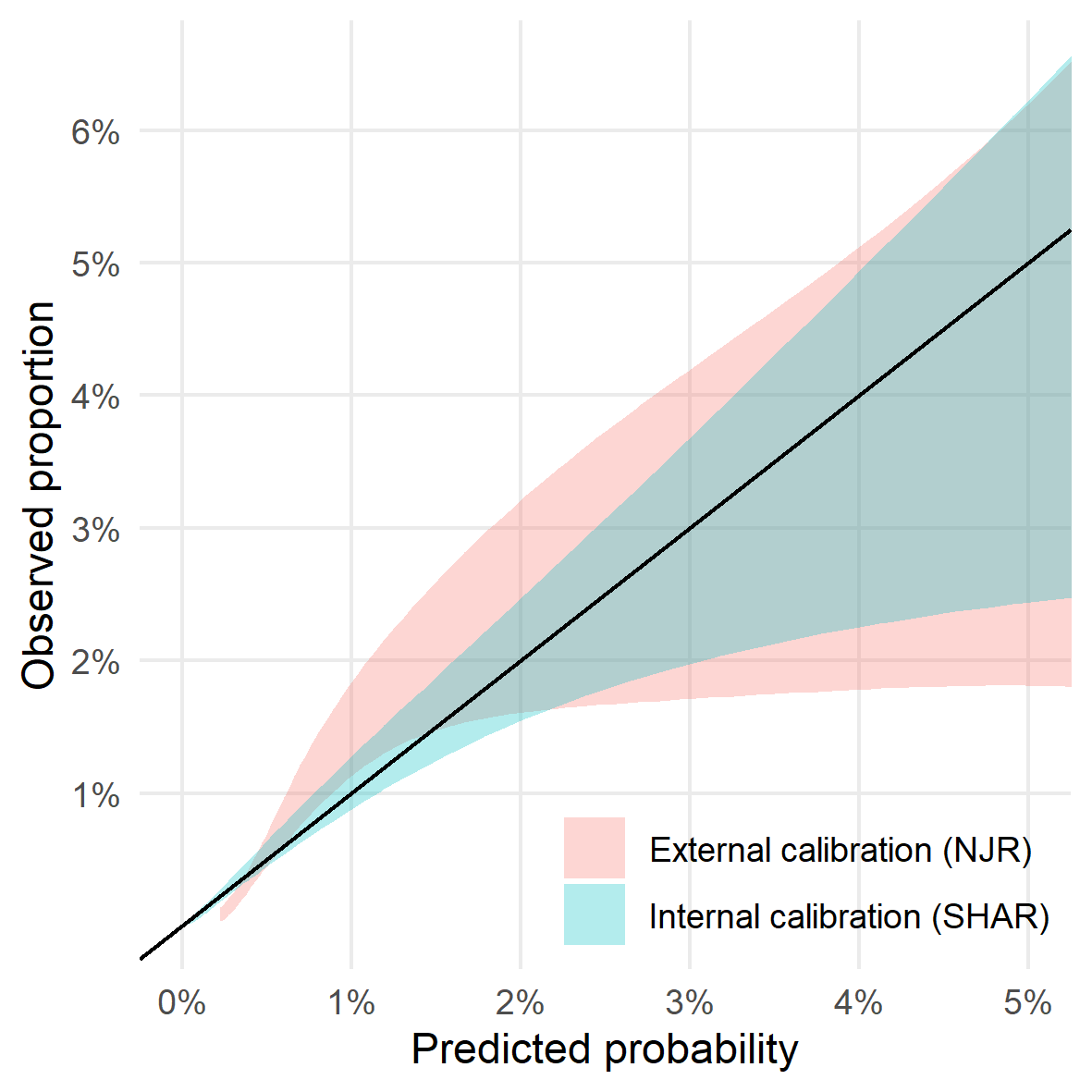


Figure 5: This figure illustrates calibrartion between observed proportions and predicted probabilities with 95 % confidence intervals. Deviations above the diagonal line means that the model underpredicts the probability of death within 90 days. Such deviation was statistically significant for the external validation for predictions below 1.5%. The higher the predicted probability gets, the more they over-estimate the observed proportions of deaths.

##### PAGE BREAK

Table 1: Categorization of co-morbdities based on diagnostic groups according to Charlson and Elixhauser. (CNS = central nervous system.)

|  |  |  |
| --- | --- | --- |
| Comorbidities by groups | Charlson | Elixhauser |
| AIDS/HIV | Aids/hiv | Aids/hiv |
| Anemia |  | Blood loss anemia, Deficiency anemia |
| Arrhythmia |  | Cardiac arrhythmias |
| Arterial hypertension |  | Hypertension uncomplicated, Hypertension complicated |
| Cancer | Malignancy, Metastatic solid tumor | Lymphoma, Metastatic cancer, Solid tumor |
| CNS disease | Dementia, Hemiplegia or paraplegia | Depression, Paralysis, Other neurological disorders, Psychoses |
| Coagulopathy |  | Coagulopathy |
| Diabetes | Diabetes without complication, Diabetes complication | Diabetes uncomplicated, Diabetes complicated |
| Drug alcohol abuse |  | Alcohol abuse, Drug abuse |
| Fluid electrolyte disorders |  | Fluid electrolyte disorders |
| Heart condition | Congestive heart failure | Congestive heart failure, Valvular disease |
| Heart infarct | Myocardial infarction |  |
| Hypothyroidism |  | Hypothyroidism |
| Kidney disease | Renal disease | Renal failure |
| Liver disease | Mild liver disease, Moderate or severe liver disease | Liver disease |
| Lung airways disease | Chronic pulmonary disease | Chronic pulmonary disease, Pulmonary circulation disorder |
| Obesity |  | Obesity |
| Peptic ulcer | Peptic ulcer disease | Peptic ulcer disease |
| Rheumatic disease | Rheumatic disease | Rheumatoid arthritis |
| Vascular disease | Peripheral vascular disease, Cerebrovascular disease | Peripheral vascular disorder |
| Weight loss |  | Weight loss |

##### PAGE BREAK

Table 2: Characteristics of the study population. Co-morbidities recorded for at least one patient who survived 90 days, and one who did not, were modeled as potential predictors. (CNS = central nervous system.)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| what | level | Died within 90 days | Survived at least 90 days | Total |
| n |  | 175 | 52924 | 53099 |
| Age (mean (SD)) |  | 77.99 (7.89) | 72.66 (7.76) | 72.68 (7.76) |
| Sex = Female (%) |  | 77 (44.0) | 32363 (61.1) | 32440 (61.1) |
| BMI (mean (SD)) |  | 26.81 (5.18) | 27.19 (4.39) | 27.19 (4.39) |
| ASA class (%) |  |  |  |  |
|  | I | 7 ( 4.0) | 9582 (18.1) | 9589 (18.1) |
|  | II | 86 (49.1) | 33795 (63.9) | 33881 (63.8) |
|  | III | 82 (46.9) | 9547 (18.0) | 9629 (18.1) |
| Hospital (%) |  |  |  |  |
|  | University | 74 (42.3) | 24386 (46.1) | 24460 (46.1) |
|  | County | 66 (37.7) | 16441 (31.1) | 16507 (31.1) |
|  | Rural | 19 (10.9) | 9921 (18.7) | 9940 (18.7) |
|  | Private | 16 ( 9.1) | 2176 ( 4.1) | 2192 ( 4.1) |
| Education (%) |  |  |  |  |
|  | Low | 30 (17.1) | 11615 (21.9) | 11645 (21.9) |
|  | Middle | 82 (46.9) | 20522 (38.8) | 20604 (38.8) |
|  | High | 63 (36.0) | 20787 (39.3) | 20850 (39.3) |
| Civil status (%) |  |  |  |  |
|  | Married | 83 (47.4) | 29353 (55.5) | 29436 (55.4) |
|  | Single | 38 (21.7) | 12850 (24.3) | 12888 (24.3) |
|  | Widow/widower | 54 (30.9) | 10721 (20.3) | 10775 (20.3) |
| Charlson (%) |  |  |  |  |
|  | 0 | 78 (44.6) | 39178 (74.0) | 39256 (73.9) |
|  | 1 | 41 (23.4) | 8076 (15.3) | 8117 (15.3) |
|  | 2 | 25 (14.3) | 3737 ( 7.1) | 3762 ( 7.1) |
|  | 3 | 12 ( 6.9) | 1164 ( 2.2) | 1176 ( 2.2) |
|  | 4+ | 19 (10.9) | 769 ( 1.5) | 788 ( 1.5) |
| Elixhauser (%) |  |  |  |  |
|  | 0 | 56 (32.0) | 27717 (52.4) | 27773 (52.3) |
|  | 1 | 46 (26.3) | 13720 (25.9) | 13766 (25.9) |
|  | 2 | 31 (17.7) | 7208 (13.6) | 7239 (13.6) |
|  | 3+ | 42 (24.0) | 4279 ( 8.1) | 4321 ( 8.1) |
| AIDS/HIV (%) |  | 0 | 5 ( <0.1) | 5 ( 0.0) |
| Anemia (%) |  | 4 ( 2.3) | 412 ( 0.8) | 416 ( 0.8) |
| Arrhythmia (%) |  | 32 (18.3) | 4473 ( 8.5) | 4505 ( 8.5) |
| Arterial hypertension (%) |  | 70 (4<0.1) | 16607 (31.4) | 16677 (31.4) |
| Cancer (%) |  | 26 (14.9) | 2689 ( 5.1) | 2715 ( 5.1) |
| CNS disease (%) |  | 14 ( 8.0) | 1668 ( 3.2) | 1682 ( 3.2) |
| Coagulopathy (%) |  | 0 | 192 ( 0.4) | 192 ( 0.4) |
| Diabetes (%) |  | 26 (14.9) | 4051 ( 7.7) | 4077 ( 7.7) |
| Drug alcohol abuse (%) |  | 1 ( 0.6) | 222 ( 0.4) | 223 ( 0.4) |
| Fluid electrolyte disorders (%) |  | 0 | 304 ( 0.6) | 304 ( 0.6) |
| Heart condition (%) |  | 31 (17.7) | 2608 ( 4.9) | 2639 ( 5.0) |
| Heart infarct (%) |  | 23 (13.1) | 2163 ( 4.1) | 2186 ( 4.1) |
| Hypothyroidism (%) |  | 7 ( 4.0) | 1784 ( 3.4) | 1791 ( 3.4) |
| Kidney disease (%) |  | 14 ( 8.0) | 537 ( 1.0) | 551 ( 1.0) |
| Liver disease (%) |  | 0 | 207 ( 0.4) | 207 ( 0.4) |
| Lung airways disease (%) |  | 18 (10.3) | 2860 ( 5.4) | 2878 ( 5.4) |
| Obesity (%) |  | 7 ( 4.0) | 993 ( 1.9) | 1000 ( 1.9) |
| Peptiulcer (%) |  | 2 ( 1.1) | 339 ( 0.6) | 341 ( 0.6) |
| Rheumatidisease (%) |  | 10 ( 5.7) | 1912 ( 3.6) | 1922 ( 3.6) |
| Vascular disease (%) |  | 15 ( 8.6) | 1671 ( 3.2) | 1686 ( 3.2) |
| Weight loss (%) |  | 0 | 35 ( 0.1) | 35 ( 0.1) |

##### PAGE BREAK

Table 3: Variables selected by the bootstrap ranking procedure, and therefore kept in the main model. Variables chosen at least 33 times out of 100 were kept in the reduced model as well. ASA class II and III were kept in both models, since those are usually distinguished anyway. To simplify the models further would be possible by lumping class I and II together. Variables chosen more frequently are likely more important predictors. This is not necessarily true however since one of several strongly correlated variables might be chosen spuriously. (CNS = central nervous system.)

|  |  |
| --- | --- |
| variable | n |
| Cancer | 100 |
| CNS disease | 100 |
| Kidney disease | 100 |
| ASA class III | 100 |
| Age | 50 |
| Obesity | 49 |
| Male sex | 38 |
| Heart condition | 20 |
| Anemia | 9 |
| ASA class II | 8 |
| Heart infarct | 6 |

##### PAGE BREAK

Table 4: Estimated coefficients (beta) and odds ratios (OR) with 95% confidence intervals for the reduced model.

|  |  |  |  |
| --- | --- | --- | --- |
| term | beta | OR 95 % CI | p |
| (intercept) | -13.28 |  |  |
| Cancer | 0.71 | 2.04 (1.30-3.08) | <0.01 |
| CNS disease | 0.90 | 2.45 (1.35-4.11) | <0.01 |
| Kidney disease | 1.33 | 3.79 (2.04-6.52) | <0.01 |
| ASA class II | 0.88 | 2.42 (1.20-5.78) | 0.03 |
| ASA class III | 1.66 | 5.27 (2.56-12.77) | <0.01 |
| Age | 0.08 | 1.08 (1.06-1.11) | <0.01 |
| Obesity | 0.77 | 2.17 (0.91-4.36) | 0.05 |
| Male sex | 0.61 | 1.85 (1.36-2.52) | <0.01 |

##### PAGE BREAK

# Bibliography

1 Hunt LP, Ben-Shlomo Y, Clark EM *et al.* 90-day mortality after 409 096 total hip replacements for osteoarthritis, from the National Joint Registry for England and Wales: A retrospective analysis. *The Lancet* 2013; **382**: 1097–104.

2 Garland A, Gordon M, Garellick G, Kärrholm J, Sköldenberg O, Hailer NP. Risk of early mortality after cemented compared with cementless total hip arthroplasty. *The Bone & Joint Journal* 2017; **99-B**: 37–43.

3 Inacio MCS, Pratt NL, Roughead EE, Graves SE. Using medications for prediction of revision after total joint arthroplasty. *The Journal of arthroplasty* 2015; **30**: 2061–70.

4 Gordon M, Stark A, Sköldenberg OG, Kärrholm J, Garellick G. The influence of comorbidity scores on re-operations following primary total hip replacement: Comparison and validation of three comorbidity measures. *The bone & joint journal* 2013; **95-B**: 1184–91.

5 Hofstede SN, Gademan MGJ, Vlieland TPMV, Nelissen RGHH, de Mheen PJM-v. Preoperative predictors for outcomes after total hip replacement in patients with osteoarthritis: A systematic review. *BMC musculoskeletal disorders* 2016; **17**: 212.

6 Bülow E, Rolfson O, Cnudde P, Rogmark C, Garellick G, Nemes S. Comorbidity does not predict long-term mortality after total hip arthroplasty. *Acta Orthopaedica* 2017; **88**: 1–6.

7 Brusselaers N, Lagergren J. The charlson comorbidity index in registry-based research. *Methods of Information in Medicine* 2017; **56**: 401–6.

8 Walraven CV, Austin PC, Jennings A *et al.* A modification of the elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med* 2009; **47**: 626–33.

9 Bozic KJ, Ong K, Lau E *et al.* Estimating risk in medicare patients with THA: An electronic risk calculator for periprosthetic joint infection and mortality. *Clinical Orthopaedics and Related Research* 2013; **471**: 574–83.

10 Bozic KJ. Orthopaedic healthcare worldwide: Shared medical decision making in orthopaedics. *Clinical Orthopaedics and Related Research* 2013; **471**: 1412–4.

11 Harris AHS, Kuo AC, Bozic KJ *et al.* American joint replacement registry risk calculator does not predict 90-day mortality in veterans undergoing total joint replacement. *Clinical Orthopaedics and Related Research* 2018; **476**: 1869–75.

12 Harris AH, Kuo AC, Bowe T, Gupta S, Nordin D, Giori NJ. Prediction models for 30-Day mortality and complications after total knee and hip arthroplasties for veteran health administration patients with osteoarthritis. *The Journal of Arthroplasty* 2018; **33**: 1539–45.

13 Inacio MCCS, Pratt NLL, Roughead EEE, Graves SEE. Evaluation of three co-morbidity measures to predict mortality in patients undergoing total joint arthroplasty. *Osteoarthritis and cartilage* 2016; **24**: 1718–26.

14 Price A, Smith J, Dakin H *et al.* The Arthroplasty Candidacy Help Engine tool to select candidates for hip and knee replacement surgery: Development and economic modelling. *Health Technology Assessment* 2019; **23**: 1–216.

15 Harris AHS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? *Clinical Orthopaedics and Related Research* 2019; **477**: 452–60.

16 Manning DW, Edelstein AI, Alvi HM. Risk prediction tools for hip and knee arthroplasty. *The Journal of the American Academy of Orthopaedic Surgeons* 2016; **24**: 19–27.

17 Bülow E, Nemes S, Rolfson O. Are the first or the second hips of staged bilateral THAs more similar to unilateral procedures? A study from the swedish hip arthroplasty register. *Clinical Orthopaedics and Related Research* 2020;: 1.

18 Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research. *European Journal of Epidemiology* 2009; **24**: 659–67.

19 Cnudde P, Rolfson O, Nemes S *et al.* Linking Swedish health data registers to establish a research database and a shared decision-making tool in hip replacement. *BMC Musculoskeletal Disorders* 2016; **17**: 414.

20 Kärrholm J, Mohaddes M, Odin D, Vinblad J, Rogmark C, Rolfson O. Svenska höftprotesregistret Årsrapport 2017. 2018.

21 Ludvigsson JF, Svedberg P, Olén O, Bruze G, Neovius M. The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *European Journal of Epidemiology* 2019; **34**: 423–37.

22 Ludvigsson JF, Andersson E, Ekbom A *et al.* External review and validation of the Swedish national inpatient register. *BMC Public Health* 2011; **11**: 450.

23 Brophy S, John G, Evans E, Lyons RA. Methodological issues in the identification of hip fractures using routine hospital data: A database study. *Osteoporosis International* 2006; **17**: 405–9.

24 Burns EM, Rigby E, Mamidanna R *et al.* Systematic review of discharge coding accuracy. *Journal of Public Health* 2012; **34**: 138–48.

25 Sinha S, Peach G, Poloniecki JD, Thompson MM, Holt PJ. Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomesSystematic review and recommendations for reporting. *European Journal of Public Health* 2013; **23**: 86–92.

26 Simmonds SJ, Syddall HE, Walsh B *et al.* Understanding NHS hospital admissions in england: Linkage of hospital episode statistics to the hertfordshire cohort study. *Age and Ageing* 2014; **43**: 653–60.

27 Thorn JC, Turner E, Hounsome L *et al.* Validation of the hospital episode statistics outpatient dataset in england. *PharmacoEconomics* 2016; **34**: 161–8.

28 Sayers A, Ben-Shlomo Y, Blom AW, Steele F. Probabilistic record linkage. *International Journal of Epidemiology* 2016; **45**: 954–64.

29 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases* 1987; **40**: 373–83.

30 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology* 1992; **45**: 613–9.

31 Quan H, Sundararajan V, Halfon P *et al.* Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical care* 2005; **43**: 1130–9.

32 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Medical care* 1998; **36**: 8–27.

33 Guo P, Zeng F, Hu X *et al.* Improved variable selection algorithm using a LASSO-Type penalty, with an application to assessing hepatitis b infection relevant factors in community residents. *PLOS ONE* 2015; **10**: e0134151.

34 Baranowski R, Chen Y, Fryzlewicz P. Ranking-based variable selection for high-dimensional data. *Statistica Sinica* 2020. DOI:[10.5705/ss.202017.0139](https://doi.org/10.5705/ss.202017.0139).

35 Austin PC, Tu JV. Bootstrap methods for developing predictive models. *The American Statistician* 2004; **58**: 131–7.

36 Ripley B, Venables WN. Modern applied statistics with S. Springer, New York, 2002.

37 Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters* 2006; **27**: 861–74.

38 Miller ME, Hui SL, Tierney WM. Validation techniques for logistic regression models. *Statistics in Medicine* 1991; **10**: 1213–26.

39 Nattino G, Finazzi S, Bertolini G. A new test and graphical tool to assess the goodness of fit of logistic regression models. *Statistics in Medicine* 2016; **35**: 709–20.

40 Steyerberg EW, Borsboom GJJM, van Houwelingen HC, Eijkemans MJC, Habbema JDF. Validation and updating of predictive logistic regression models: A study on sample size and shrinkage. *Statistics in Medicine* 2004; **23**: 2567–86.

41 Breiman L. Statistical modeling: The two cultures. *Statistical Science* 2001; **16**: 199–215.

42 Haynes SR, Lawler PG. An assessment of the consistency of ASA physical status classification allocation. *Anaesthesia* 1995; **50**: 195–9.

43 Whitmore RG, Stephen JH, Vernick C *et al.* ASA grade and Charlson Comorbidity Index of spinal surgery patients: Correlation with complications and societal costs. *The spine journal : official journal of the North American Spine Society* 2014; **14**: 31–8.

44 Kork F, Balzer F, Krannich A, Weiss B, Wernecke K-D, Spies C. Association of comorbidities with postoperative in-hospital mortality: A retrospective cohort study. *Medicine* 2015; **94**: e576.

45 Mood C. Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review* 2010; **26**: 67–82.

46 Söderman P, Malchau H, Herberts P, Johnell O. Are the findings in the swedish national total hip arthroplasty register valid? A comparison between the swedish national total hip arthroplasty register, the national discharge register, and the national death register. *The Journal of arthroplasty* 2000; **15**: 884–9.

47 Söderman P, Malchau H, Herberts P, Zügner R, Regnér H, Garellick G. Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the swedish national total hip arthroplasty register. *Acta orthopaedica Scandinavica* 2001; **72**: 113–9.

48 Spencer SA, Davies MP. Hospital episode statistics: Improving the quality and value of hospital data: A national internet e-survey of hospital consultants. *BMJ Open* 2012; **2**. DOI:[10.1136/bmjopen-2012-001651](https://doi.org/10.1136/bmjopen-2012-001651).

49 Williams JG, Mann RY. Hospital episode statistics: Time for clinicians to get involved? *Clinical medicine (London, England)* 2002; **2**: 34–7.

50 Brennan L, Watson M, Klaber R, Charles T. The importance of knowing context of hospital episode statistics when reconfiguring the NHS. *BMJ (Online)* 2012; **344**. DOI:[10.1136/bmj.e2432](https://doi.org/10.1136/bmj.e2432).

51 Smith AJ, Dieppe P, Porter M, Blom AW. Risk of cancer in first seven years after metal-on-metal hip replacement compared with other bearings and general population: Linkage study between the National Joint Registry of England and Wales and hospital episode statistics. *BMJ (Online)* 2012; **344**. DOI:[10.1136/bmj.e2383](https://doi.org/10.1136/bmj.e2383).

52 Mouchti S, Whitehouse MR, Sayers A, Hunt LP, MacGregor A, Blom AW. The association of body mass index with risk of long-term revision and 90-Day mortality following primary total hip replacement. *The Journal of Bone and Joint Surgery* 2018; **100**: 2140–52.

53 Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-reported weight and height: Findings from the NHANES & NHIS 2001-2006. *BMC Public Health* 2009; **9**: 1–10.