Prediction of 90-day mortality after Total Hip Arthroplasty: a simplified and extrenaly validated model based on Swedish and English register data

Anne Garland1,2,3, Erik Bulow2,4, Adrian Sayers5,6, Szilard Nemes2,4, Ola Rolfson2,4, Nils P. Hailer1,2

2019-03-12

# 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. …
6. …

**Correspondence:** AG: [anne.l.garland@gmail.com](mailto:anne.l.garland@gmail.com), Phone: +46 498 26 80 00, Postal: Orthopeadic Department, Visby Hospital, St Göransgatan 10, 621 84 Visby, Sweden  
EB: [erik.bulow@registercentrum.se](mailto:erik.bulow@registercentrum.se)  
SN: [szilard.nemes@registercentrum.se](mailto:szilard.nemes@registercentrum.se)  
NH: [nils.hailer@surgsci.uu.se](mailto:nils.hailer@surgsci.uu.se)

# Abstract

**OBJECTIVE:** Early mortality after total hip arthroplasty (THA) is associated with comorbidity. Diagnosis-based tools such as the Charlson or Elixhauser Comorbidity indices, or the prescription-based RxRisk V score, have been used to quantify this. Those instruments are however complex and requires the use of extensive data sets not routinely available in most doctor-patient settings. We investigated if a simplified model, could substitute those complex indices to predict early mortality after THA.

**PATIENTS AND METHODS:** We studied 49235 patients with THA due to primary osteoarthritis 2008 - 2015 from the Swedish Hip Arthroplasty Register, linked to national the population register, the National Patient Register and the longitudinal integration database for health insurance and labour market studies from Statistics Sweden. We used stepwise forward and backward logistic regression in combination with bootstrap averaging (bagging), to develop a prediction model for patient deaths within 90 days after surgery. Predictive power was assessed by the area under the reciever operating characteristic curve (AUC). The final model was applied to British data for external validation.

**RESLUTS:** The unadjusted cumulative 90-day survival was 99.68 (95 % CI: 99.62 - 99.73). Best predictive performance for 90-day mortality was found for a model combining Age, ASA, Sex, or the precense of cancer, heart infarct, kidney disease or gout, during the year before surgery (AUC = 0.78). This mdoel was superior to the established but complex comorbidity measures RxRiskV Score (0.61), the Charlson comorbidity index (AUC = 0.67), and the Elixhauser comorbidity score (AUC = 0.64).

**CONCLUSION:** We found a relatively simple prediction model of 90 day mortality after total hip arthroplasty. This model requires less data and is easier to calculate compared to previously well-known comorbidity indices.

# Introduction

The presence of pre-surgery comorbidity is associated with poorer outcome after the insertion of total hip arthroplasty (THA). Patients with increased comorbidity an increased risk of early postoperative mortality and revision (Inacio et al. 2015; Gordon et al. 2013), as well as inferior patient-reported outcomes (Gordon et al. 2013; Hofstede et al. 2016). In research settings, comorbidity is commonly measured using multi-facetted diagnosis- or prescription-based coding algorithms (Bozic et al. 2013). Inacio et al. (2016) studied the ability of three commonly used coding algorithms to predict mortality after THA and total knee arthroplasty: the Charlson (CCI) and Elixhauser (ECI) Comorbidity Indices, as well as RxRiskV. These comorbidity measures are quite complex to estimate and are based on the availability of extensive datasets including in- and outpatient data on ICD-codes, or detailed information on drug prescriptions prior to surgery. Oftentimes, such datasets can only be created by linking several population-based registries, raising both ethical and practical concerns. Also, each of the comorbidity index exist in numerous versions (Sundararajan et al. 2004; Deyo, Cherkin, and Ciol 1992; Quan et al. 2011; Cleves, Sanchez, and Draheim 1997; Walraven et al. 2009). Interpretation and comparison between different studies is therefore difficult. Comorbidity data have also been used in several universal and arthroplasty-specific risk prediction tools to make risk profiles for individual patients. In the context of trauma, prediction tools are common, and it has been possible to reduce the number of variables without losing predictive power (Gerdin et al. 2016). No model has so far been broadly accepted for elective THA however (Manning, Edelstein, and Alvi 2016). An easily applicable tool with few dimensions is thus needed, both in research and in clinical practice. We aimed to find such a model to predict the risk of 90-day postoperative mortality after THA.

# Patients and Methods

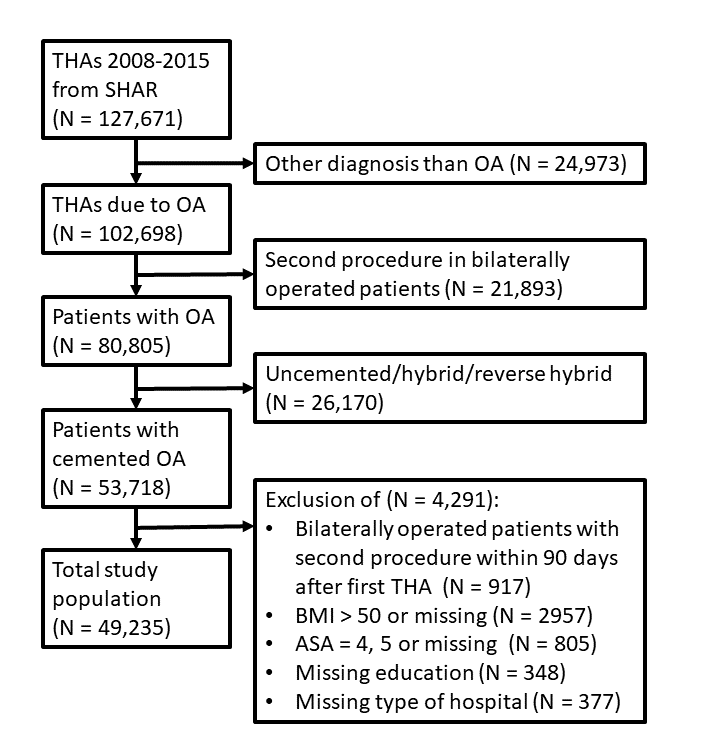


Figure 1 Flowchart depicting inclusion criteria and number of patients included in the development phase of the model.

Patients recorded in the Swedish Hip Arthroplasty Register (SHAR) with cemented primary osteoarthritis 2008 - 2015 were included in the development phase of the study (Figure 1). Only the first THA was accounted for in bilaterally operated patients and patients were excluded if a second THA was inserted within 90 days after the first. Hence, the ninety-day mortality were not influenced by any subsequent THA surgery. Follow-up started on the day of surgery and ended at death, emigration, or December 31st 2015, whichever came first.

Data linkage, based on the unique identity numbers assigned to all Swedish residents (Ludvigsson et al. 2009), were used to collect data from a variaty of sources, as previsouly described by Cnudde et al. (2016).

Age, sex, body mass index (BMI), ASA class, type of hospital (university/county/rural/private) and year of surgery were collected from SHAR, with a completeness of 96-98% (Kärrholm et al. 2018). Data on education level (low/middle/high) and civil status, were collected from the longitudinal integration database for health insurance and labour market studies from Statistics Sweden (Ludvigsson et al. 2019). The Swedish National Patient Register was used for comorbidity data during one year before surgery. The register contains all relevant diagnoses coded by ICD-10, as well as admissions and discharge dates for in- and outpatient visits in all private and public hospitals (Ludvigsson et al. 2011). We used medical data from the national prescription register where all prescriptions are classified by the Anatomic Therapeutic Chemical classification system (ATC). Death dates were linked from the national population register.

Comorbidity was recognized by individual ICD-10 and ATC codes grouped into 17 categories according to CCI (Charlson et al. 1987; Deyo, Cherkin, and Ciol 1992; Quan et al. 2005), 31 categories according to ECI and 42 categories according to the RxRiskV (Fishman et al. 2003; Sloan et al. 2003) classification. RxRiskV medications prescribed for inflammatory pain was excluded since those are also sold over-the-counter in grocery stores, yielding questionable completeness in the prescription register. Patients with no hospital visits or no prescriptions were assumed to have no comorbidity.

Some comorbidities were identified by more than one classification and some distinct comorbidities were closely related (such as hypertension with and without complications, or abuse of either drugs or alcohol). We identified 16 broader categories in addition to 5 standalone ECI classes and 14 RxRiskV classes that were kept unchanged. Groups were merged according to clinical relevance as to be recognized in a patient-doctor meeting without access to external register data (see supplementary data).

## Statistics

We used the Kaplan-Meier estimator to assess unadjusted cumulative survival.

Further analysis was based on logistic regression since no censoring occoured within the 90 day study period.

Univariable models were applied to all predictors individually to estimate odds ratios for descriptive purposes.

All potential predictors were then partioned into four groups where there were assumed to be no colinearity among variables: 1) the 17 combined comorbidities, 2) the 5 ECI comorbidities, 3) the 14 RxRiskV comorbidities, and 4) age, sex, BMI, ASA, type of hospital, year of surgery, education and civil status.

1,000 bootstrap samples were drawn from the initial data set. Early mortality is very rare for patients with elective THA however. Downsampling stratified on the outcome was therefore applied to each sample to include only as many survivals as deaths. This gave us balanced samples and managable computations. (King and Zeng 2001; Courvoisier et al. 2011; Austin and Steyerberg 2017).

A step-wise linear regression procedure based on the Aikake Information Criteria (AIC) was applied to each bootstrap sample to identify important predictors from each of the four groups of variables (Sauerbrei 1999; Zellner, Keller, and Zellner 2004; Lukacs, Burnham, and Anderson 2010). Variables that were kept in the final model for at least 75 % of all bootstrap samples were selected as candidates for further consideration.

A full main effects model were made to include all previously selected candidate predictors. Two modified versions were considered as well, where age was modeled by resttricted cubic spline with either 2 or 3 knots. Three simplified models with age and sex only, were used for comparison, as well as four univariable regression models with ASA, CCI, ECI and RxRiskV alone.

We then took another 1,000 downsampled bootsrtap samples to estimate coefficients for each model. Model averaging based on the information criteria was used for the full main effects model and its variants. Hence, an automated model selection procedure evaluating all possible subsets of the included main effects, yielding averaged coeffients estimates based on variable importance.

Each model, now with their coefficient estimates fixed, were then evaluated by their predictive power based on their area under the reciever operation characteristics (ROC) curve (AUC). Point estimates as well as 95 % confidence intervals were estimated by emperical bootstrap, once again with 1,000 resamples but this time without downsampling. Sensitivity and specificity from each bootstrap sample were subjected to a local polynomial regression fit (loess) to illustrate an averaged reciever operation characteristics (ROC) curve for each model.

We used R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) with significant packages tidyverse, tidymodels, furrr, pROC, MuMIn and rFSA (Lambert et al. 2018).

## Ethical approval

Ethical approval for this study was obtained from the Regional Ethical Review Board in Gothenburg (271-14 and 360-13).

# Results

There were 49235 patients (Figure 1), 35 - 98 years old, whereof 61 % were female. 158 (0.32 %) patients died within 90 days and no one was censored before that. The unadjusted cummulative 90-day survival was 99.68 (95 % CI: 99.62 - 99.73).

Some characteristics of the study population are presented in Table 1. 94.33 % of all patients had at least one pre-surgery comorbidity according to RxRiskV, but only 25.64 % and 46.84 % according to ECI and CCI. The proportion of patients with ASA class 3 was 17.66 %. Most individual comorbidities were more common among patients who died.

Table 1 Baseline demographic information on the study population 2008-2012. CCI/ECI = Charlson/Elixhauser Comorbidity Index. Individual comorbidities without prefixes were based on combinations of different comorbidity classifications. Prefix ECI and Rx(RiskV) indicate classification specific comorbidities.

|  |  |  |  |
| --- | --- | --- | --- |
| variable | level | alive | dead |
| N |  | 49077 | 158 |
| Age (mean (SD)) |  | 72.35 (7.78) | 78.27 (7.80) |
| Sex = Female (%) |  | 29960 (61.0) | 68 (43.0) |
| BMI (mean (SD)) |  | 27.18 (4.38) | 26.73 (5.19) |
| ASA (%) |  |  |  |
|  | 1 | 9163 (18.7) | 7 ( 4.4) |
|  | 2 | 31292 (63.8) | 76 (48.1) |
|  | 3 | 8622 (17.6) | 75 (47.5) |
| Hospital (%) |  |  |  |
|  | university | 22668 (46.2) | 66 (41.8) |
|  | county | 15125 (30.8) | 61 (38.6) |
|  | rural | 9333 (19.0) | 17 (10.8) |
|  | private | 1951 ( 4.0) | 14 ( 8.9) |
| Education (%) |  |  |  |
|  | low | 10807 (22.0) | 26 (16.5) |
|  | middle | 18905 (38.5) | 75 (47.5) |
|  | high | 19365 (39.5) | 57 (36.1) |
| Civil status (%) |  |  |  |
|  | married | 27363 (55.8) | 73 (46.2) |
|  | single | 12010 (24.5) | 36 (22.8) |
|  | widow/widower | 9704 (19.8) | 49 (31.0) |
| CCI (%) |  |  |  |
|  | 0 | 36538 (74.5) | 71 (44.9) |
|  | 1 | 7366 (15.0) | 36 (22.8) |
|  | 2 | 3425 ( 7.0) | 21 (13.3) |
|  | 3 | 1044 ( 2.1) | 11 ( 7.0) |
|  | 4+ | 704 ( 1.4) | 19 (12.0) |
| ECI (%) |  |  |  |
|  | 0 | 26120 (53.2) | 51 (32.3) |
|  | 1 | 12599 (25.7) | 40 (25.3) |
|  | 2 | 6497 (13.2) | 28 (17.7) |
|  | 3+ | 3861 ( 7.9) | 39 (24.7) |
| RxRiskV (%) |  |  |  |
|  | 0 | 2786 ( 5.7) | 6 ( 3.8) |
|  | 1 | 5573 (11.4) | 10 ( 6.3) |
|  | 2 | 6880 (14.0) | 12 ( 7.6) |
|  | 3 | 7137 (14.5) | 14 ( 8.9) |
|  | 4 | 6846 (13.9) | 20 (12.7) |
|  | 5 | 6227 (12.7) | 29 (18.4) |
|  | 6 | 5216 (10.6) | 23 (14.6) |
|  | 7+ | 8412 (17.1) | 44 (27.8) |
| Heart condition (%) |  | 21755 (44.3) | 99 (62.7) |
| Heart infarct (%) |  | 21599 (44.0) | 105 (66.5) |
| Arrythmia (%) |  | 4194 ( 8.5) | 32 (20.3) |
| Hypertoni (%) |  | 24090 (49.1) | 105 (66.5) |
| Kidney disease (%) |  | 585 ( 1.2) | 15 ( 9.5) |
| Diabetes (%) |  | 5065 (10.3) | 30 (19.0) |
| Vascular disease (%) |  | 16223 (33.1) | 77 (48.7) |
| Cancer (%) |  | 2594 ( 5.3) | 25 (15.8) |
| Aids hiv (%) |  | 13 ( 0.0) | 0 ( 0.0) |
| Liver disease (%) |  | 3077 ( 6.3) | 13 ( 8.2) |
| Lung airways disease (%) |  | 5747 (11.7) | 29 (18.4) |
| Drug alcohol abuse (%) |  | 278 ( 0.6) | 2 ( 1.3) |
| Cns (%) |  | 9785 (19.9) | 40 (25.3) |
| Reuma (%) |  | 1735 ( 3.5) | 8 ( 5.1) |
| Anemia (%) |  | 354 ( 0.7) | 4 ( 2.5) |
| Peptiulcer (%) |  | 11916 (24.3) | 39 (24.7) |
| ECI hypothyroidism (%) |  | 1601 ( 3.3) | 6 ( 3.8) |
| ECI coagulopathy (%) |  | 167 ( 0.3) | 0 ( 0.0) |
| ECI obesity (%) |  | 902 ( 1.8) | 6 ( 3.8) |
| ECI weight loss (%) |  | 39 ( 0.1) | 0 ( 0.0) |
| ECI fluid electrolyte disorders (%) |  | 275 ( 0.6) | 0 ( 0.0) |
| Rx allergies (%) |  | 4639 ( 9.5) | 13 ( 8.2) |
| Rx benign prostate hypertrophy (%) |  | 1477 ( 3.0) | 8 ( 5.1) |
| Rx glaucoma (%) |  | 2684 ( 5.5) | 8 ( 5.1) |
| Rx gout (%) |  | 1268 ( 2.6) | 15 ( 9.5) |
| Rx hyperkalaemia (%) |  | 46 ( 0.1) | 0 ( 0.0) |
| Rx hyperlipidemia (%) |  | 12962 (26.4) | 42 (26.6) |
| Rx hyperthyroidism (%) |  | 4601 ( 9.4) | 14 ( 8.9) |
| Rx malnutrition (%) |  | 1 ( 0.0) | 0 ( 0.0) |
| Rx osteoporosis pagets (%) |  | 2229 ( 4.5) | 9 ( 5.7) |
| Rx pain (%) |  | 17024 (34.7) | 59 (37.3) |
| Rx pancreatiinsufficiency (%) |  | 131 ( 0.3) | 1 ( 0.6) |
| Rx psoriasis (%) |  | 379 ( 0.8) | 2 ( 1.3) |
| Rx steroid responsive diseases (%) |  | 5033 (10.3) | 26 (16.5) |
| Rx transplant (%) |  | 85 ( 0.2) | 0 ( 0.0) |

Odds ratios from univariable logistic regresison models are presented in Table 2. Higher age, ASA, CCI, ECI and RxRiskV, as well as being male or widow(er), or having cancer, gout, kidney disease or heart infarct, increase the risk of death within 90 days after surgery. Private hospitals tend to operate healthier patients and the odds of death decrease slightly by year of surgery.

Table 2 Unadjusted odds ratios (OR) for 90-day mortality with 95 % confidence intervals (CI). Individual comorbiditities are only presented if selected in at least 75 % of all bootstrap samples (selected). Variables included in the final model are marked with ‘FM’ in model column.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| variable | level | OR | CI | selected (%) | model |
| Age |  | 1.11 | (1.09-1.14) | 100 | FM |
| Sex | female | 0.48 | (0.35-0.66) | 99 | FM |
| BMI |  | 0.98 | (0.94-1.01) | 21 |  |
| ASA | 2 | 3.18 | (1.47-6.9) | 63 | FM |
|  | 3 | 11.39 | (5.25-24.71) | 100 | FM |
| Hospital | county | 1.39 | (0.98-1.96) | 19 |  |
|  | rural | 0.63 | (0.37-1.07) | 20 |  |
|  | private | 2.46 | (1.38-4.4) | 39 |  |
| Year |  | 0.92 | (0.86-0.99) | 52 |  |
| Education | middle | 1.65 | (1.05-2.58) | 21 |  |
|  | high | 1.22 | (0.77-1.95) | 19 |  |
| Civil status | single | 1.12 | (0.75-1.68) | 26 |  |
|  | widow/widower | 1.89 | (1.32-2.72) | 53 |  |
| Rx |  | 1.21 | (1.12-1.31) |  |  |
| ECI |  | 1.69 | (1.47-1.93) |  |  |
| CCI |  | 1.86 | (1.66-2.08) |  |  |
| Cancer |  | 3.37 | (2.19-5.17) | 92 | FM |
| Heart infarct |  | 2.52 | (1.81-3.51) | 79 | FM |
| Kidney disease |  | 8.69 | (5.08-14.89) | 81 | FM |
| Rx gout |  | 3.95 | (2.32-6.75) | 90 | FM |

ASA class was the only variable included in the final models by all stage-wise regression procedures from each of the 1,000 bootstrap samples (100 % in column “selected” in Table 2). Most potential predictors were rarely selected and only Age, ASA, Sex, or the precense of cancer, heart infarct, kidney disease or gout, during the year before surgery, where selected at least 75 % of the time (Table 2).

Table 3 Area Under the Curve (AUC) as a measure of predictive power for the full model compared to simple models with standard comorbidity measures, as well as a reference model with age and sex. Age was included as either a main effect, or in the form of restricted cubic splines (rcs) with either 2 or 3 knots.

|  |  |  |  |
| --- | --- | --- | --- |
| dimension | model | age | AUC |
| Multivariable | age + sex | main | 0.74 (0.68-0.79) |
|  |  | rcs 2 | 0.59 (0.46-0.64) |
|  |  | rcs 3 | 0.59 (0.46-0.64) |
|  | full | main | 0.78 (0.72-0.82) |
|  |  | rcs 2 | 0.73 (0.68-0.78) |
|  |  | rcs 3 | 0.74 (0.69-0.79) |
|  | full-cancer |  | 0.78 (0.73-0.83) |
|  | simpl |  | 0.77 (0.72-0.82) |
| Univariable | ASA |  | 0.34 (0.29-0.73) |
|  | CCI |  | 0.67 (0.36-0.72) |
|  | ECI |  | 0.64 (0.35-0.70) |
|  | RxRiskV |  | 0.61 (0.55-0.67) |

A model with those covariates, the “full” model, had an out-of-sample AUC of 0.78. The predictive power did not improve with age modeled by restricted cubic splines (Table 3 and Figure 2). The full model led to better predictions compared to a simplified model with only age and sex, or any of the conventional comorbidity indices (CCI, ECI and RxRiskV). The full model is further described in Table 4.

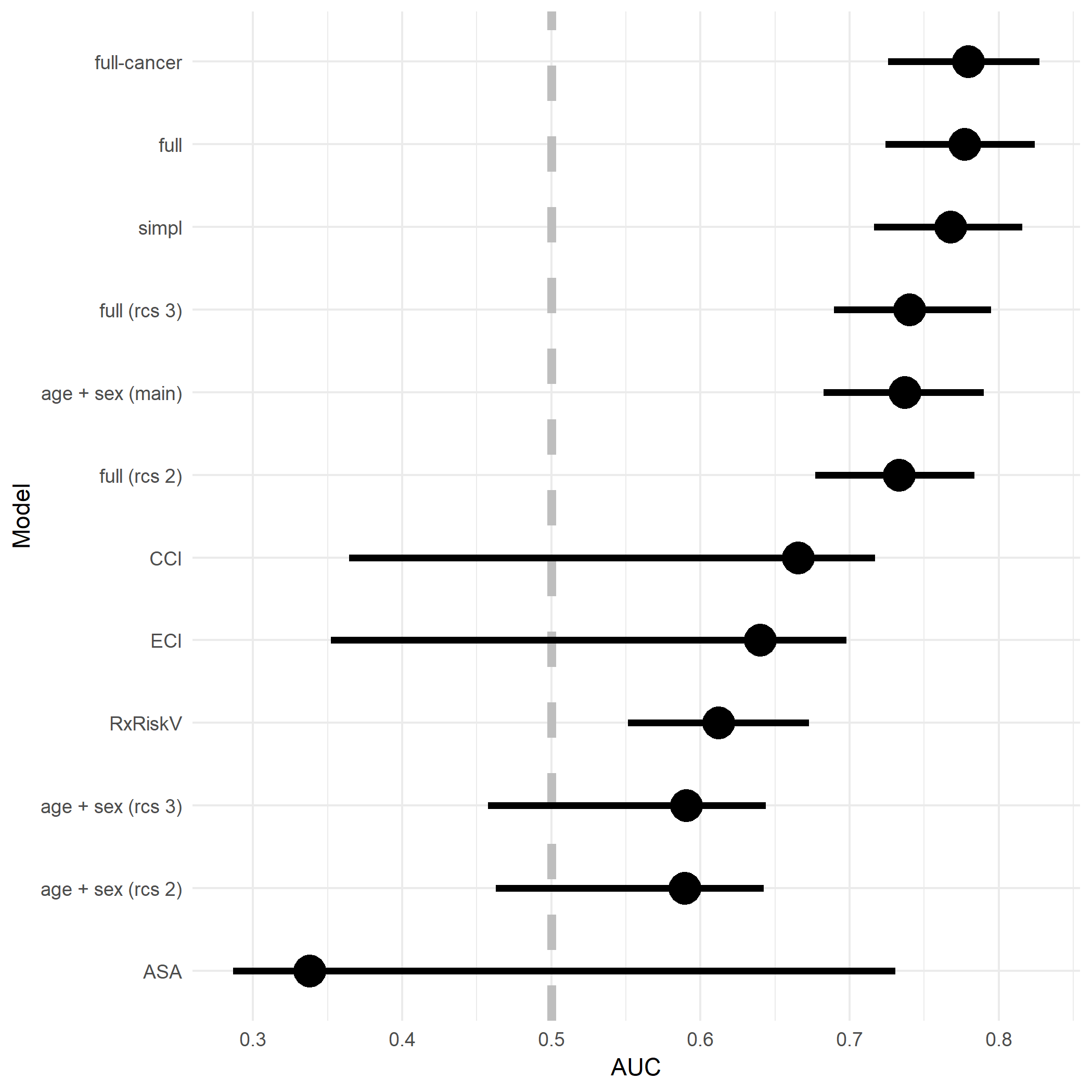


Figure 2 Area Under the Curve (AUC) as a measure of predictive power for the full model compared to simple models with standard comorbidity measures, as well as a reference model with age and sex. Age was included as either a main effect, or in the form of restricted cubic splines (rcs) with 2 or 3 knots. Confidence intervals based on 1,000 non-paramteric bootstrap replicates.

Table 4 Estimated odds ratios for variables in the full model based on full averages from 1,000 bootstrap samples.

|  |  |  |  |
| --- | --- | --- | --- |
| variables | OR | CI | p |
| Age | 1.1e+00 | 1.05-1.12 | < 0.001 |
| ASA 2 | 1.4e+00 | 0.51-3.98 | 0.5 |
| ASA 3 | 2.1e+00 | 0.47-9.05 | 0.33 |
| cancer | 1.5e+00 | 0.59-3.65 | 0.4 |
| Female | 5.5e-01 | 0.29-1.06 | 0.07 |
| heart infarct | 1.5e+00 | 0.70-3.24 | 0.29 |
| kidney disease | 9.4e+00 | 0.95-92.90 | 0.06 |
| Rx gout | 7.9e+06 | 0.00-Inf | 0.99 |

ROC curves for all models compared to the full model are displayed in Figure 3. Each models ability separate survivors from death are further illustrated in Figure 4.

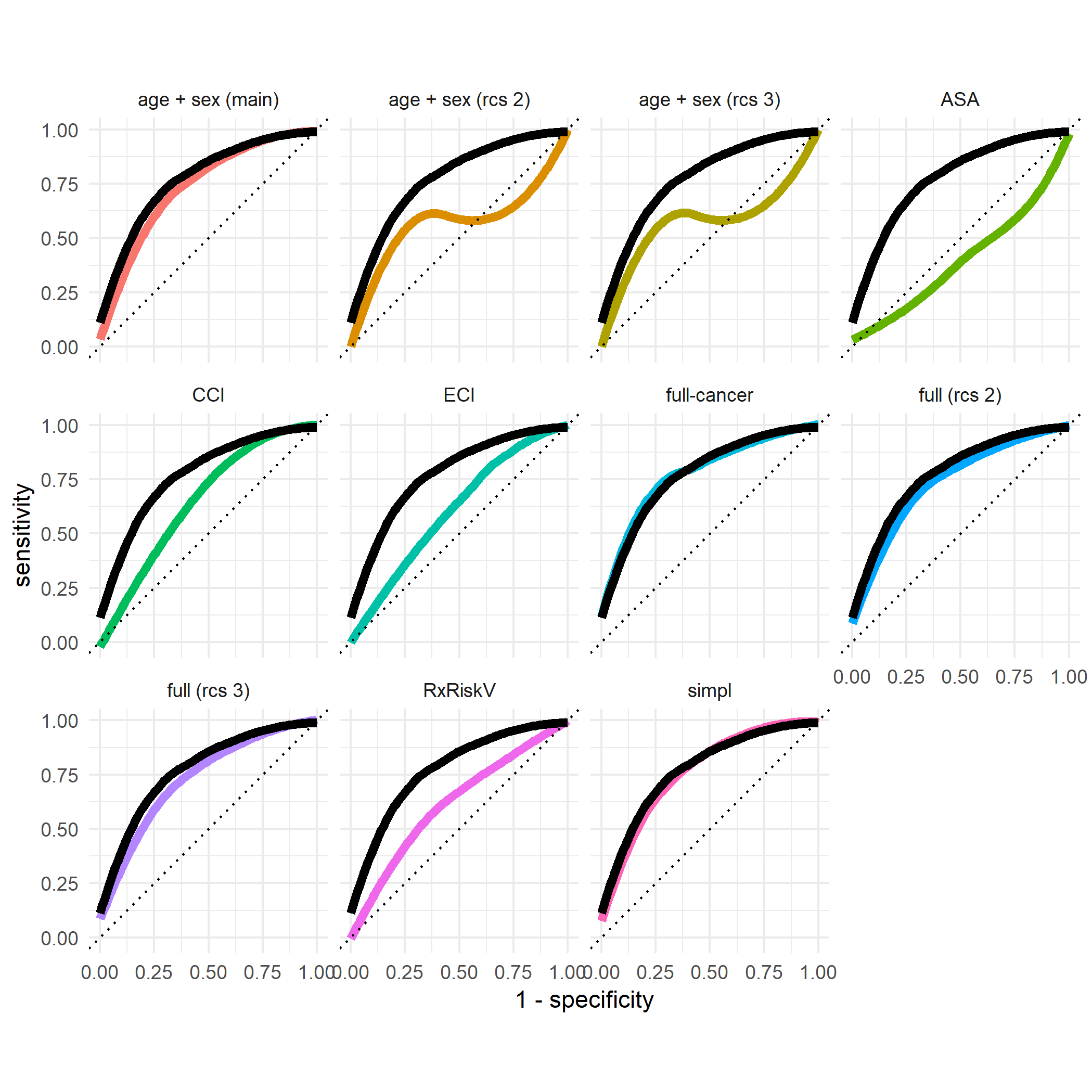


Figure 3 ROC curves for each model compared to the full main effects model.

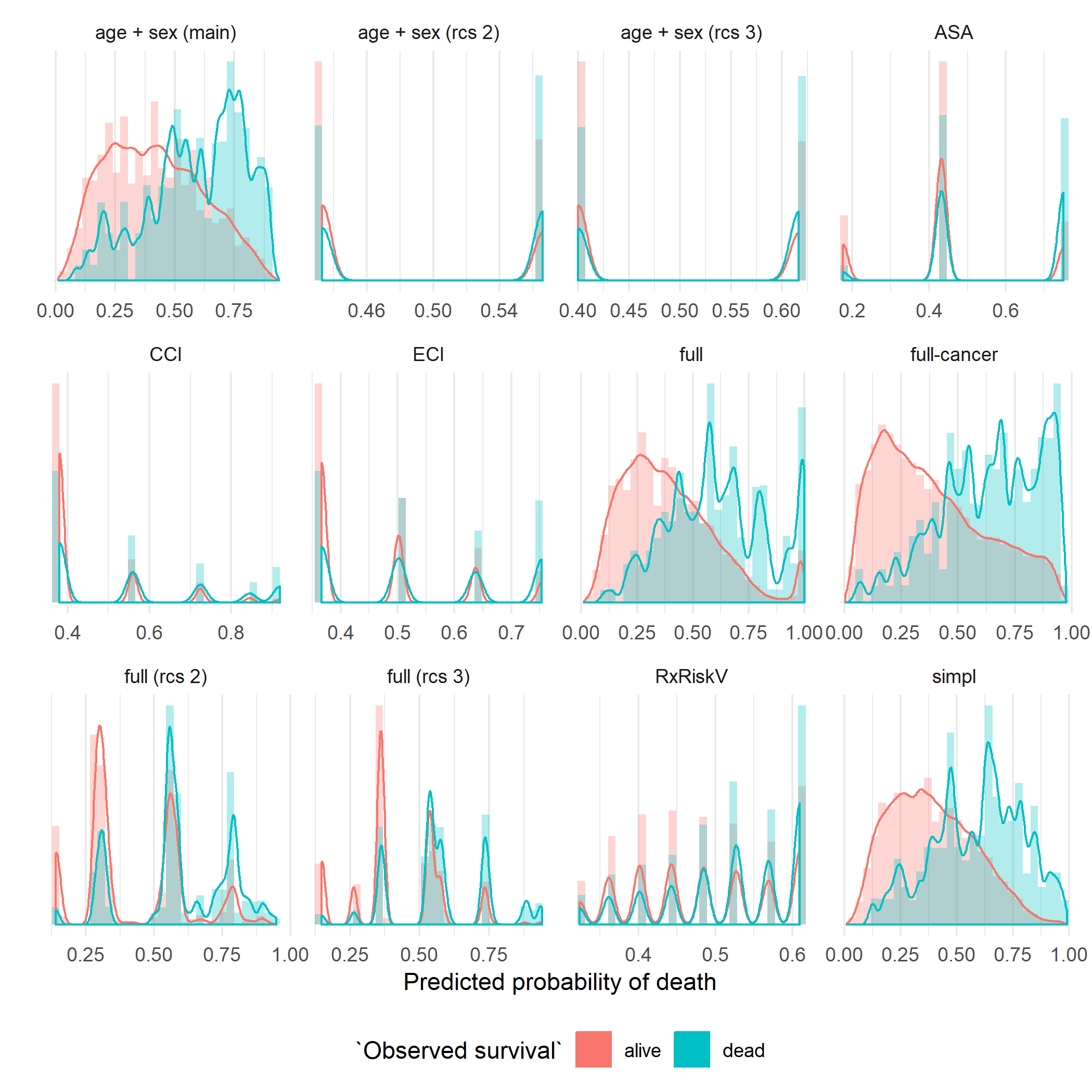


Figure 4 Separation

# Discussion

We found that a univariable main effects logistic regression model with Age, ASA, Sex, or the precense of cancer, heart infarct, kidney disease or gout, during the year before surgery was superior to predict death within 90 days after insertion of elective THA. Some covariates are not statistically significant themselves (4) but are still relevant due to unobserved heteregeneity (Mood 2010).

It is well known that male sex is associated with earlier deaths. It is also well known that the remaining life span will decrease with incresed age. It is however not obvious that this relation must be striclty linear. We therefore allowed for a more flexible relation based on restricted cubic splines, but found that a linear relationship was indeed most sufficient.

ASA class was the only covariate chosen by all stepwise regression procedures. This might not be surprising due to its definition. It is interesting to note however that a univariable model with ASA grade alone, yielded the lowest predictive power of all models compared. An out-of-sample point estimate of its AUC was < 0.5 (4 and 3, meaning that predictions based solely on ASA would be worse than random. This shows the important of validation, due to the risk of overfitting, as well as the importance of multivariable adjustment. ASA class is routinely assessed pre-operatively in most developed countries. It is however known to have a high degree of internal variability (Haynes and Lawler 1995). It has previously been compared to the CCI, but not with respect to mortality after THA (Whitmore et al. 2014; Kork et al. 2015). It has been argued that a prescription-based measure would be more reliable than a diagnosed based measure (Iezzoni 1997; Johnson et al. 2006). We did confirm that estimation of AUC were more stable for the RxRiskV index compared to CCI and ECI (as seen by the width of the confidence intervals in figure 2), although point estimates of the latter was marginaly higher. This reflects that both coverage and completeness from the prescription register is better compared to the national patient regiseter. The first hold records of all medications prescribed, the second only contains information that the doctor deemed relevant at the patient encounter.

# Further discussion from Garland not yet adjusted to new models

A set of easily accessible data is a better predictor of early mortality after major surgery than complex coding algorithms In this nationwide cohort study we intended to compare the performance of a set of easily accessible data that are routinely collected in daily clinical practice with complex comorbidity coding algorithms (ie CCI, Elixhauser Score and RxRiskV). The best predictive strength was found for a relatively simple model including age, gender, presence of cardiac infarction or renal disease during the last 12 months prior to THA surgery, and ASA grade (c=0.81). Thios simple model was also better than the above mentioned comorbidity measures at predicting one-year mortality (Supplementary Figure 1 and Supplementary Table 3). Comorbidities are known to influence the outcome after THA (Inacio et al. 2015; Gordon et al. 2013; Hofstede et al. 2016). In order to assess the effect of comorbidity on early mortality after THA different coding algorithms have been proposed in research settings. The coding algorithms are complex, and hence they demand a merge of information on ICD-ocdes and medication prescriptions from several data sources. These coding algorithms are not used in clinical settings since the administrative burden associated with identifying some 30 ICD- or ATC-codes for every patient is not realistic. Thus, comorbidity measures based on patient administrative data are only accessible to researchers, but - even then -observational study designs are hampered by the usual limitations such as incompleteness and inaccuracy of coding (Bozic et al. 2013). In this present study we found that the prescription-based RxRiskV Score performed better than the diagnosis-based comorbidity comorbidity measures CCI and Elixhauser Score in predicting 90-day mortality. The original CCI was somewhat better than the Elixhauser Score in predicting 90-day mortality. This differs from earlier findings by Inacio et al. where the RxRiskV did not perform as well as the CCI and the Elixhauser Score, and where the c- statistics were generally higher than in our study (Inacio et al. 2016). Such dissimilarities could be explained by the facts that our study population was younger, included more women, and that we only included diagnoses and prescriptions registered one year prior to surgery. Overall, the predictive strength of all investigated diagnose- or prescription-based comorbidity measures was better than the included dimensions investigated separately. To put it differently, in terms of predicting mortality, each comorbidity measure was an improvement over the separate items included in each measure. We also found that the ASA classification was better at predicting both 90-day and one-year mortality than the more complex coding algorithms, with a c- statistic of 0.70. The ASA classification has been repeatedly compared to the CCI, but no consensus as to which one is superior has been reached, and, to our knowledge, such comparisons have not been performed on THA population (Whitmore et al. 2014; Kork et al. 2015). Individuals with an ASA grade of 4 to 6 were excluded from our study since those categories describe severe disease, moribund and brain-dead individuals, and it can be questioned whether the classification is correct, and - if it is - whether these patients should ever have received a THA. We thus excluded this very small group. Obesity is generally known to be associated with a higher risk of morbidity and all-cause mortality (Must and McKeown 2000). However, previous studies on primary THA cohorts have not indicated a higher risk of mortality in obese patients, a result that is confirmed in our study (Supplementary Table 1) (Wallace et al. 2014). An explanation could be that obese patients selected for THA are comparably healthy.

Risk prediction may be useful in the process of patient selection prior to surgery, in the preoperative risk management including a review of current medications, and in perioperative anaesthesia management. A number of risk prediction tools of various complexity for adverse outcomes after total joint replacements have been introduced but none has been broadly accepted (Manning, Edelstein, and Alvi 2016). In the context of trauma surgery outcome prediction tools are common, and it is seems possible to reduce the number of items without losing predictive power (Gerdin et al. 2016). Our results indicate that the risk of early postoperative mortality after THA could be assessed by a relatively simple prediction model.

A strength of this study is its nationwide design with a large cohort of primary THA patients with a reasonable number of events, rendering estimation relatively precise. Our sources of data are highly valid, and the proportion of missing data in our cohort was low (Söderman et al. 2000, 2001; Garellick et al. 2014; Ludvigsson et al. 2011). Limitations to this study are the potential biases at different levels that are commonly associated with observational data, and the risk of coding errors as expected when dealing with patient administrative data. Selection bias is also an issue in this study, since patients who died on the table during attempted THA surgery may not have been reported to the SHAR, but, judging from clinical experience, such events are extremely uncommon in a population of patients scheduled for elective THA surgery for osteoarthritis. The much higher expected frequency of fatal events during THA surgery on hip fracture patients was one of the main reasons to exclude this group from the present analyses. It is important to distinguish between explanatory observational research and attempts at predicting individual events such as early mortality after surgical interventions. The combination of parameters in the best-performing model described in the present study may serve as a predictor of mortality on an individual level, but the described combination of parameters would need to be validated in a different sample of individuals. Since this has not yet been done we cannot extrapolate our findings to prediction models in a clinical setting, but aim at performing such additional studies Our results indicate that in research on mortality after a very common surgical intervention a less complex comorbidity measure consisting of easily accessible data that are routinely collected in daily clinical practice is superior to some of the commonly used diagnose- or prescription-based coding algorithms. It would be interesting to evaluate the ability of our novel set of parameters to predict adverse events and revision rates.

Our results derived from a nationwide cohort study indicate that a less data demanding comorbidity measure, the combination of age, gender, presence of heart infarction or renal disease and ASA grade, serves is better at predicting early postoperative mortality after THA than comorbidity measures based on more complex coding algorithms.

# Contribution of authors

AG and NH initiated the study and managed the ethical review board application. EB, AS, SN, and AG performed the statistical analyses. AG and EB drafted the manuscript. All athours edited and finalized the manuscript.

# Bibliography

Austin, Peter C, and Ewout W Steyerberg. 2017. “Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models.” *Statistical Methods in Medical Research* 26 (2). SAGE Publications: 796–808. <https://doi.org/10.1177/0962280214558972>.

Bozic, Kevin J, Ravi K Bashyal, Shawn G Anthony, Vanessa Chiu, Brandon Shulman, and Harry E Rubash. 2013. “Is administratively coded comorbidity and complication data in total joint arthroplasty valid?” *Clinical Orthopaedics and Related Research* 471 (1): 201–5. <https://doi.org/10.1007/s11999-012-2352-1>.

Charlson, Mary E., Peter Pompei, Kathy L. Ales, and C. Ronald MacKenzie. 1987. “A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation.” *Journal of Chronic Diseases* 40 (5): 373–83. <https://doi.org/10.1016/0021-9681(87)90171-8>.

Cleves, M A, N Sanchez, and M Draheim. 1997. “Evaluation of two competing methods for calculating Charlson’s comorbidity index when analyzing short-term mortality using administrative data.” *Journal of Clinical Epidemiology* 50 (8): 903–8. <http://www.ncbi.nlm.nih.gov/pubmed/9291875>.

Cnudde, Peter, Ola Rolfson, Szilard Nemes, Johan Kärrholm, Clas Rehnberg, Cecilia Rogmark, John Timperley, and Göran Garellick. 2016. “Linking Swedish health data registers to establish a research database and a shared decision-making tool in hip replacement.” *BMC Musculoskeletal Disorders* 17 (1). BMC Musculoskeletal Disorders: 414. <https://doi.org/10.1186/s12891-016-1262-x>.

Courvoisier, Delphine S., Christophe Combescure, Thomas Agoritsas, Angèle Gayet-Ageron, and Thomas V. Perneger. 2011. “Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure.” *Journal of Clinical Epidemiology* 64 (9). Pergamon: 993–1000. <https://doi.org/10.1016/J.JCLINEPI.2010.11.012>.

Deyo, Richard A., Daniel C. Cherkin, and Marcia A. Ciol. 1992. “Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases.” *Journal of Clinical Epidemiology* 45 (6). Pergamon: 613–19. <https://doi.org/10.1016/0895-4356(92)90133-8>.

Fishman, Paul A, Michael J Goodman, Mark C Hornbrook, Richard T Meenan, Donald J Bachman, and Maureen C O’Keeffe Rosetti. 2003. “Risk adjustment using automated ambulatory pharmacy data: the RxRisk model.” *Medical Care* 41 (1): 84–99. <https://doi.org/10.1097/01.MLR.0000039830.19812.29>.

Garellick, Göran, Johan Kärrholm, Hans Lindahl, Henrik Malchau, Cecilia Rogmark, and Ola Rolfson. 2014. “Svenska Höftprotesregistret: Årsrapport 2014.” Gothenburg, Sweden. <https://doi.org/10.18158/SJhEVmKcb>.

Gerdin, Martin, Nobhojit Roy, Monty Khajanchi, Vineet Kumar, Li Felländer-Tsai, Max Petzold, Göran Tomson, Johan von Schreeb, and Towards Improved Trauma Care Outcomes in India (TITCO). 2016. “Validation of a novel prediction model for early mortality in adult trauma patients in three public university hospitals in urban India.” *BMC Emergency Medicine* 16 (February): 15. <https://doi.org/10.1186/s12873-016-0079-0>.

Gordon, M, A Stark, O G Sköldenberg, J Kärrholm, and G Garellick. 2013. “The influence of comorbidity scores on re-operations following primary total hip replacement: comparison and validation of three comorbidity measures.” *The Bone & Joint Journal* 95-B (9). British Editorial Society of Bone; Joint Surgery: 1184–91. <https://doi.org/10.1302/0301-620X.95B9.31006>.

Haynes, S R, and P G Lawler. 1995. “An assessment of the consistency of ASA physical status classification allocation.” *Anaesthesia* 50 (3): 195–9. <http://www.ncbi.nlm.nih.gov/pubmed/7717481>.

Hofstede, Stefanie N, Maaike G J Gademan, Thea P M Vliet Vlieland, Rob G H H Nelissen, and Perla J Marang-van de Mheen. 2016. “Preoperative predictors for outcomes after total hip replacement in patients with osteoarthritis: a systematic review.” *BMC Musculoskeletal Disorders* 17: 212. <https://doi.org/10.1186/s12891-016-1070-3>.

Iezzoni, L I. 1997. “Assessing quality using administrative data.” *Annals of Internal Medicine* 127 (8 Pt 2): 666–74. <http://www.ncbi.nlm.nih.gov/pubmed/9382378>.

Inacio, Maria C S, Nicole L Pratt, Elizabeth E Roughead, and Stephen E Graves. 2015. “Using Medications for Prediction of Revision after Total Joint Arthroplasty.” *The Journal of Arthroplasty* 30 (12): 2061–70. <https://doi.org/10.1016/j.arth.2015.06.009>.

Inacio, M.C.S. C S, N.L. L Pratt, E.E. E Roughead, and S.E. E Graves. 2016. “Evaluation of three co-morbidity measures to predict mortality in patients undergoing total joint arthroplasty.” *Osteoarthritis and Cartilage* 24 (10). W.B. Saunders: 1718–26. <https://doi.org/10.1016/j.joca.2016.05.006>.

Johnson, Michael L, Hashem B El-Serag, Tung Thomas Tran, Christine Hartman, Peter Richardson, and Neena S Abraham. 2006. “Adapting the Rx-Risk-V for mortality prediction in outpatient populations.” *Medical Care* 44 (8): 793–7. <https://doi.org/10.1097/01.mlr.0000218804.41758.ef>.

Kärrholm, Johan, Maziar Mohaddes, Daniel Odin, Johanna Vinblad, Cecilia Rogmark, and Ola Rolfson. 2018. “Svenska Höftprotesregistret Årsrapport 2017.” Gothenburg, Sweden. <https://registercentrum.blob.core.windows.net/shpr/r/-rsrapport-2016-SJirXXUsb.pdf>.

King, Gary, and Langche Zeng. 2001. “Logistic Regression in Rare Events Data.” *Political Analysis* 9 (02). Cambridge University Press: 137–63. <https://doi.org/10.1093/oxfordjournals.pan.a004868>.

Kork, Felix, Felix Balzer, Alexander Krannich, Björn Weiss, Klaus-Dieter Wernecke, and Claudia Spies. 2015. “Association of comorbidities with postoperative in-hospital mortality: a retrospective cohort study.” *Medicine* 94 (8): e576. <https://doi.org/10.1097/MD.0000000000000576>.

Lambert, Joshua, Liyu Gong, Corrine F Elliott, Katherine Thompson, and Arnold Stromberg. 2018. “rFSA: An R Package for Finding Best Subsets and Interactions.” *The R Journal* 10 (2): 295–308. <https://journal.r-project.org/archive/2018/RJ-2018-059/RJ-2018-059.pdf>.

Ludvigsson, Jonas F, Eva Andersson, Anders Ekbom, Maria Feychting, Jeong-Lim Kim, Christina Reuterwall, Mona Heurgren, and Petra Otterblad Olausson. 2011. “External review and validation of the Swedish national inpatient register.” *BMC Public Health* 11 (1): 450. <https://doi.org/10.1186/1471-2458-11-450>.

Ludvigsson, Jonas F., Petra Otterblad-Olausson, Birgitta U. Pettersson, and Anders Ekbom. 2009. “The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research.” *European Journal of Epidemiology* 24 (11): 659–67. <https://doi.org/10.1007/s10654-009-9350-y>.

Ludvigsson, Jonas F., Pia Svedberg, Ola Olén, Gustaf Bruze, and Martin Neovius. 2019. “The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research.” *European Journal of Epidemiology* 34 (4). Springer Netherlands: 423–37. <https://doi.org/10.1007/s10654-019-00511-8>.

Lukacs, Paul M., Kenneth P. Burnham, and David R. Anderson. 2010. “Model selection bias and Freedman’s paradox.” *Annals of the Institute of Statistical Mathematics* 62 (1). Springer-Verlag: 117–25. <https://doi.org/10.1007/s10463-009-0234-4>.

Manning, David W, Adam I Edelstein, and Hasham M Alvi. 2016. “Risk Prediction Tools for Hip and Knee Arthroplasty.” *The Journal of the American Academy of Orthopaedic Surgeons* 24 (1): 19–27. <https://doi.org/10.5435/JAAOS-D-15-00072>.

Mood, C. 2010. “Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It.” *European Sociological Review* 26 (1). Oxford University Press: 67–82. <https://doi.org/10.1093/esr/jcp006>.

Must, Aviva, and Nicola M McKeown. 2000. *The Disease Burden Associated with Overweight and Obesity*. <http://www.ncbi.nlm.nih.gov/pubmed/25905320>.

Quan, Hude, Bing Li, Chantal M. Couris, Kiyohide Fushimi, Patrick Graham, Phil Hider, Jean Marie Januel, and Vijaya Sundararajan. 2011. “Updating and validating the charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries.” *American Journal of Epidemiology* 173 (6): 676–82. <https://doi.org/10.1093/aje/kwq433>.

Quan, Hude, Vijaya Sundararajan, Patricia Halfon, Andrew Fong, Bernard Burnand, Jean-Christophe Luthi, L Duncan Saunders, Cynthia a Beck, Thomas E Feasby, and William a Ghali. 2005. “Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data.” *Medical Care* 43 (11): 1130–9. <https://doi.org/10.1097/01.mlr.0000182534.19832.83>.

Sauerbrei, Willi. 1999. “The Use of Resampling Methods to Simplify Regression Models in Medical Statistics.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 48 (3). John Wiley & Sons, Ltd (10.1111): 313–29. <https://doi.org/10.1111/1467-9876.00155>.

Sloan, Kevin L, Anne E Sales, Chuan-Fen Liu, Paul Fishman, Paul Nichol, Norman T Suzuki, and Nancy D Sharp. 2003. “Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument.” *Medical Care* 41 (6): 761–74. <https://doi.org/10.1097/01.MLR.0000064641.84967.B7>.

Söderman, P, H Malchau, P Herberts, and O Johnell. 2000. “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* 15 (7): 884–9. <http://www.ncbi.nlm.nih.gov/pubmed/11061449>.

Söderman, P, H Malchau, P Herberts, R Zügner, H Regnér, and G Garellick. 2001. “Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the Swedish National Total Hip Arthroplasty Register.” *Acta Orthopaedica Scandinavica* 72 (2): 113–9. <https://doi.org/10.1080/000164701317323345>.

Sundararajan, Vijaya, Toni Henderson, Catherine Perry, Amanda Muggivan, Hude Quan, and William A Ghali. 2004. “New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality.” *Journal of Clinical Epidemiology* 57 (12): 1288–94. <https://doi.org/10.1016/j.jclinepi.2004.03.012>.

Wallace, G, A Judge, D Prieto-Alhambra, F de Vries, N K Arden, and C Cooper. 2014. “The effect of body mass index on the risk of post-operative complications during the 6 months following total hip replacement or total knee replacement surgery.” *Osteoarthritis and Cartilage* 22 (7): 918–27. <https://doi.org/10.1016/j.joca.2014.04.013>.

Walraven, Carl Van, Peter C Austin, Alison Jennings, Hude Quan, J Forster Alan, Carl Van Walraven, Peter C Austin, and Alison Jennings. 2009. “A Modification of the Elixhauser Comorbidity Measures into a Point System for Hospital Death Using Administrative Data.” *Med* 47 (6): 626–33.

Whitmore, Robert G, James H Stephen, Coleen Vernick, Peter G Campbell, Sanjay Yadla, George M Ghobrial, Mitchell G Maltenfort, and John K Ratliff. 2014. “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* 14 (1): 31–38. <https://doi.org/10.1016/j.spinee.2013.03.011>.

Zellner, Dietmar, Frieder Keller, and Günter E Zellner. 2004. “Variable Selection in Logistic Regression Models.” *Www.dekker.com COMMUNICATIONS IN STATISTICS Simulation and Computation Õ* 33 (3): 787–805. <https://doi.org/10.1081/SAC-200033363>.