## Supplementary tables

Table S3: Categorization of comorbidities from the Charlson and Elixhauser comorbidities. (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 |
| Myocardial infarction | Myocardial infarction |  |
| Hypothyroidism |  | Hypothyroidism |
| Kidney disease | Renal disease | Renal failure |
| Liver disease | Mild liver disease, Moderate or severe liver disease | Liver disease |
| Lung and 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 |

Table S4: Codes identifying PJI if recorded in the Swedish and Danish National Patient Registers within 90 days after THA.

|  |  |
| --- | --- |
| classification | codes |
| ICD-10 | M000, M000F, M001, M002, M002F, M008, M008F, M009, M009F, M860F, M861F, M866, M866F, T814, T845, T845F, T845X, T847, T847F |
| NOMESCO | NFS09, NFS19, NFS29, NFS39, NFS49, NFS59, NFS99 |

Table S5: Variables selected by the bootstrap ranking procedure. Variables selected at least 10 out of 100 times were used in the main model. Variables chosen at least 80 times were kept in the reduced model as well. (BMI = body mass index. ASA class = American Society for Anaesthesiologists classification. CNS = central nervous system.)

|  |  |
| --- | --- |
| variable | n |
| CNS disease | 100 |
| Fluid electrolyte disorders | 100 |
| Liver disease | 100 |
| ASA class: III | 100 |
| BMI: class I obesity | 100 |
| BMI: class II III obesity | 100 |
| BMI: overweight | 100 |
| Diagnosis: Avascular necrosis of the femoral head (AVN) | 100 |
| Diagnosis: Inflammatory joint disease | 100 |
| Diagnosis: Secondary osteoarthritis | 100 |
| Sex Male | 100 |
| Arrhythmia | 95 |
| Diagnosis: Sequelae after childhood hip disease | 95 |
| Lung airways disease | 93 |
| Age | 82 |
| Rheumatidisease | 68 |
| Cancer | 57 |
| Peptiulcer | 43 |
| Cemented cup | 41 |
| Hospital County | 33 |
| ASA class: II | 23 |
| Civil status widow widower | 8 |
| Hypothyroidism | 3 |
| Heart infarct | 1 |

Table S6: Model performance based on the Swedish derivation cohort. The area under the receiver operating characteristics curve (AUC) is a measure of discriminatory ability. Nagelkerke’s pseudo-coefficient of determination (R2) is an estimate of the proportion of explained variance for each model (the correlation between observed and predicted values). The univariable Charlson model had the lowest R2 and was therefore used as baseline for the fraction of new information (FNI) added by the other models.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | AUC (95 % CI) | R2 (%) | FNI (%) |
| Main model | 0.68 (0.67; 0.69) | 4.86 | 85 |
| Reduced model | 0.68 (0.66; 0.69) | 4.72 | 84 |
| ASA + BMI + Age + Sex | 0.65 (0.64; 0.66) | 3.32 | 78 |
| BMI + Age + Sex | 0.64 (0.62; 0.65) | 2.75 | 73 |
| ASA + Age + Sex | 0.62 (0.60; 0.63) | 1.94 | 62 |
| Elixhauser + Age + Sex | 0.61 (0.60; 0.62) | 1.75 | 58 |
| ASA | 0.59 (0.58; 0.60) | 1.57 | 53 |
| Elixhauser | 0.58 (0.57; 0.60) | 1.29 | 43 |
| Charlson + Age + Sex | 0.60 (0.58; 0.61) | 1.26 | 42 |
| Rx Risk | 0.58 (0.57; 0.59) | 1.03 | 29 |
| Rx Risk + Age + Sex | 0.58 (0.57; 0.59) | 1.03 | 29 |
| Charlson | 0.56 (0.55; 0.57) | 0.73 | 0 |

## 

## Model coefficients

Variable selection and coefficient estimates were based on the Swedish cohort. The reduced model was then used to predict PJI also in Denmark. The coefficient values of the reduced model were also re-fitted to the Danish cohort for comparison. We compared the estimates from the Swedish and Danish cohorts and found most values to be of similar magnitude. The only coefficient with reversed direction was for patients with a diagnose of “sequelae after childhood hip disease.” This is a rare condition, with very few observed PJI:s within 90 days. We also re-estimated the model coefficients 1,000 times based on the Swedish data. Each time, we took a random sample of the same size () as was observed in the Danish cohort. The 2.5th and 97.5th percentiles were then used to form empirical 95 % confidence intervals for each coefficient (Tab. 7 and Fig. 7). The estimated effect of “sequelae after childhood hip disease” fall outside this empirical CI. Applying multiplicity correction by Bonferroni or similar, would eliminate this significance. Hence, it seems that both cohorts are similar in respect to association between the studied covariates and the risk of PJI within 90 days of THA.

Table S7: Estimated coefficients for the reduced model based on the Swedish derivation cohort (Swedish), as well as re-estimated coefficient values based on the Danish cohort (Danish). Empirical confidence intervals (95 % CI) based on 1,000 resamples from the Swedish cohort of the same sample size (N = 18,854) as used in the Danish cohort. The Danish estimates fall within most of the CIs, indicating no support to reject the null hypothesis of no differences between the countries.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| variable | level | Swedish | Danish | 95 % CI |
| (Intercept) |  | -6.31 | -6.26 |  |
| Age |  | 0.02 | 0.02 | (0.01, 0.03) |
| Arrhythmia |  | 0.27 | 0.14 | (-0.04, 0.52) |
| ASA class | I | 0.00 | 0.00 |  |
|  | II | 0.18 | 0.38 | (-0.06, 0.45) |
|  | III | 0.44 | 0.51 | (0.12, 0.76) |
| BMI | under/normal weight | 0.00 | 0.00 |  |
|  | overweight | 0.39 | 0.42 | (0.16, 0.62) |
|  | class I obesity | 0.81 | 0.69 | (0.56, 1.08) |
|  | class II-III obesity | 1.40 | 1.55 | (1.08, 1.70) |
| CNS disease |  | 0.69 | 0.79 | (0.32, 0.99) |
| Diagnosis | Primary osteoarthritis | 0.00 | 0.00 |  |
|  | Sequelae after childhood hip disease | 0.39 | -0.52 | (-0.37, 0.89) |
|  | Avascular necrosis of the femoral head ((AVN)) | 0.58 | 0.82 | (0.00, 0.99) |
|  | Secondary osteoarthritis | 0.74 | 0.44 | (0.44, 1.00) |
|  | Inflammatory joint disease | 0.94 | 0.41 | (0.27, 1.38) |
| Fluid electrolyte disorders |  | 0.42 | 0.27 | (-0.70, 1.05) |
| Liver disease |  | 0.75 | 0.48 | (-0.20, 1.39) |
| Lung airways disease |  | 0.27 | 0.30 | (-0.07, 0.55) |
| Sex | Female | 0.00 | 0.00 |  |
|  | Male | 0.37 | 0.32 |  |

**[See separate fig file]**

Figure S7: Estimated coefficient values for the reduced model applied to the Swedish (red triangles) and Danish cohorts (red dots). Small coloured dots are empirical estimates from 1000 subsamples based on the Swedish data set, of the same sample size as used in the Danish validation cohort (outliers not shown). Black dots indicate the outer limits of 95 % empirical confidence intervals. The Danish estimates almost always fall within those CIs.