



OPEN Cause-specific mortality after spousal bereavement in a Danish register-based cohort

Mathilde Marie Brünnich Sloth^{1,2,✉}, Johannes Hruza¹, Laust Hvas Mortensen^{1,2,4}, Samir Bhatt^{1,3,5} & Alexandros Katsiferis^{1,2,5}

Spousal bereavement is associated with increased all-cause mortality, but less is known for cause-specific mortality. We estimated the average effect of bereavement on cause-specific mortality. Furthermore, we developed and evaluated the performance of prediction models estimating individualized mortality risks. This matched cohort study leveraged nationwide registry data on 223,500 married Danish individuals aged ≥ 65 years. Individuals were followed from bereavement date until death, emigration, or 3-year end of follow-up. G-estimation was used to estimate the average effect of bereavement on cause-specific mortality. Risk prediction models were developed and internally validated using logistic regression and Extreme Gradient Boosting utilizing information on sociodemographic factors and healthcare expenditures prior to bereavement. Among males, bereavement was associated with increased 3-year mortality from cardiovascular disease (CVD) (Risk Difference: 8 deaths/1000 individuals [95% confidence intervals (CI) 3;13]), digestive diseases (3 [95% CI 1;5]), psychiatric diseases/suicide (3 [95% CI 1;5]), and respiratory diseases (4 [95% CI 1;8]) and decreased 3-year mortality for dying from dementia/Parkinson's disease (-4 [95% CI -6;-2]). Among females, bereavement was only associated with increased 3-year mortality from CVD (5 [95% CI 3;7]) and from psychiatric diseases/suicide (1 [95% CI 1;2]). The prediction models developed using sociodemographic information and healthcare expenditures showed limited accuracy in identifying cause-specific mortality risk. On average, spousal bereavement slightly elevated the risk for most causes of death for males and less so for females. Despite incorporating comprehensive healthcare and sociodemographic data, accurately predicting individual cause-specific mortality remains challenging.

Keywords Spousal bereavement, Cause-specific mortality, Cohort study, Counterfactuals, Machine learning, Prediction

Spousal bereavement i.e. the loss of a spouse is a stressor often occurring in old age¹. As individuals age, they become frailer and more susceptible to stressors having an impact on their health and longevity². With an aging society, spousal bereavement becomes increasingly important to investigate. Spousal bereavement has been found to have far-reaching consequences for health. This event has been associated with various unfavorable health outcomes, such as low diet quality³, depression⁴, and mortality⁵. This underscores the urgency of understanding and addressing the impact of this life-altering experience.

Spousal bereavement has consistently been found to increase the mortality risk among spousal bereaved individuals in previous studies^{6–16}, generally with males showing a higher mortality following bereavement compared to females. Only a few studies have investigated cause-specific mortality in older adults^{13,15–17}. These studies reported higher hazard rates for spousal bereaved older adults of dying from malignant cancer, respiratory diseases, circulatory diseases, and external causes compared to non-bereaved older adults^{13,15}. However, these studies have investigated the link between spousal bereavement and cause-specific mortality with regression analysis estimating the magnitude of the association without using a causal framework. Furthermore, there is a lack of studies investigating the magnitude of the association on an absolute scale.

Recently, researchers have taken this knowledge a step further by developing prognostic prediction models that predicted all-cause mortality after spousal bereavement^{10,18}. An advantage of prognostic prediction models

¹Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.

²Data Science Lab, Statistics Denmark, Copenhagen, Denmark. ³Department of Infectious Disease Epidemiology, Imperial College London, London, UK. ⁴The ROCKWOOL Foundation, Ny Kongensgade 6, 1472, Copenhagen K, Copenhagen, Denmark. ⁵These authors jointly supervised this work: Samir Bhatt and Alexandros Katsiferis. ✉email: mathilde.sloth@sund.ku.dk

is that this method facilitates individual prediction of cause-specific risks whereas most previous research on spousal bereavement^{6–9,11–16} has used methods that estimates the average association between bereavement and all-cause mortality. Prediction models that are validated thoroughly have practical applications and can be used for resource prioritization by providing risk scores that help target those at high risk. Previous prognostic prediction models initiated at time of bereavement has found information on healthcare expenditures being particularly predictive of all-cause mortality^{10,18}. No study has to our knowledge investigated whether healthcare expenditures prior to spousal bereavement are predictive of cause-specific mortality. The current study aims to fill this gap in knowledge, building upon the aforementioned studies^{10,18}.

In a large cohort of Danish spousal bereaved older adults matched with their non-bereaved married counterparts we first aim to investigate the association between spousal bereavement and cause-specific mortality. We employed a potential outcome model¹⁹ to a causal inference analysis to ascertain whether there was a sex-specific association on average between spousal bereavement and cause-specific mortality. Secondly, we developed and internally validated prognostic models for cause-specific mortality within three years after bereavement. We used the Danish registers to include information on socioeconomic factors and health determinants including different types of healthcare expenditures as possible predictors in the models. In sum, this study aims to advance our understanding of spousal bereavement as a major life stressor and cause-specific mortality.

Methods

Study design

The original population consisted of all individuals aged 65 years and above residing in Denmark as of January 1st, 2011. Information on the setting is presented in Supplementary Text 1. We used a matched cohort study design by matching each individual experiencing spousal bereavement in 2012 with on average 20 non-bereaved controls based on a set of sociodemographic and overall health-status information at the date of bereavement (index date). Both groups were followed up from the index date until death, emigration or up to a maximum period of 3-years (administrative censoring), whichever comes first. The final matched cohort consisted of 223,500 married individuals, of which 12,940 (5.8%) experienced bereavement in 2012. The selection process and eligibility criteria of the study are described in Fig. 1.

Spousal bereavement

Spousal bereavement was defined as the loss of one's spouse. The bereavement status was included as a binary variable to determine if individuals had experienced bereavement in the year of 2012. The date of bereavement was used as time zero in the prediction analysis, i.e., the time where the prediction for an individual is being given. Information on bereavement was retrieved from the Danish Civil Registration System.

Cause-specific mortality

The outcome of the analysis was cause-specific mortality. Information on death were retrieved from the Danish Register of Causes of Death²⁰. In this study we used the Automated Classification of Medical Entities (ACME) to determine the cause of death, which is based on the causes reported on the death certificate. ACME applies the World Health Organization (WHO) rules to the International Classification of Diseases (ICD) codes and selects the underlying cause of death²¹. The following causes of death were included: dying from (1) cancer, (2) CVD, (3) diabetes, (4) dementia or Parkinson disease, (5) digestive diseases, (6) psychiatric diseases or suicide, (7) respiratory diseases and (8) other causes of death. The ICD-10 codes for each cause of death are presented in Supplementary Table 1.

Confounders and predictors

Age, sex, immigration status, affluence index, number of children and number of comorbidities measured at baseline (January 1, 2011), were included as potential confounders and predictors. This information was sourced from the Danish Civil Registration System and the National Patient Register. We additionally used longitudinal data on the healthcare expenditures of individuals 1-year prior to the index date (date of bereavement or for matched non-bereaved the date of bereavement of the matched bereaved) as another source of information potentially predictive of mortality. The Directed Acyclic Graph describing our analysis is illustrated in Supplementary Fig. 2.

Age was categorized into five-year intervals: 65–69, 70–74, 75–79, 80–84, or ≥ 85 years. Immigration status was dichotomized into 1: Danish citizens or 2: immigrants and descendants. The number of children was categorized as follows: none, 1, 2, 3, or ≥ 4 children. Comorbidities were categorized into the following groups: none, 1, 2, 3, or ≥ 4 comorbidities based on data from the creation of the National Patient Register in 1994 until the 1st January 2011. The affluence index, utilized as a measure of socioeconomic status, was derived from an individual's wealth and income based on a previous paper by Hansen and colleagues²². The healthcare expenditures entailed longitudinal information on prescription, hospital (in- and outpatient), primary care, residential care, and home care expenses. The healthcare expenditures used for this study is based on the data used in the paper by Hansen and colleagues²² and made available from Statistics Denmark. The Danish healthcare system is primarily funded by taxes, which means that the information on medical spending retrieved from the registers accounts for 97% of the spending²³. Each type of healthcare expenditure was based on weekly data in 1000s of Danish kroner (DKK) calculated as a weekly average the year prior to index date. The expenditures were categorized into four categories for the inferential analysis (the average effect of bereavement). The lowest category was total weekly average of healthcare (1) Less than 10 DKK/week and (2) 10–99 DKK/week, (3) 100–499 DKK/week, and lastly (4) the highest spending group of 500 DKK/week or more (10 DKK = 1,33 €). For the prediction modelling workflow, the healthcare expenditures were included separately as continuous predictors.

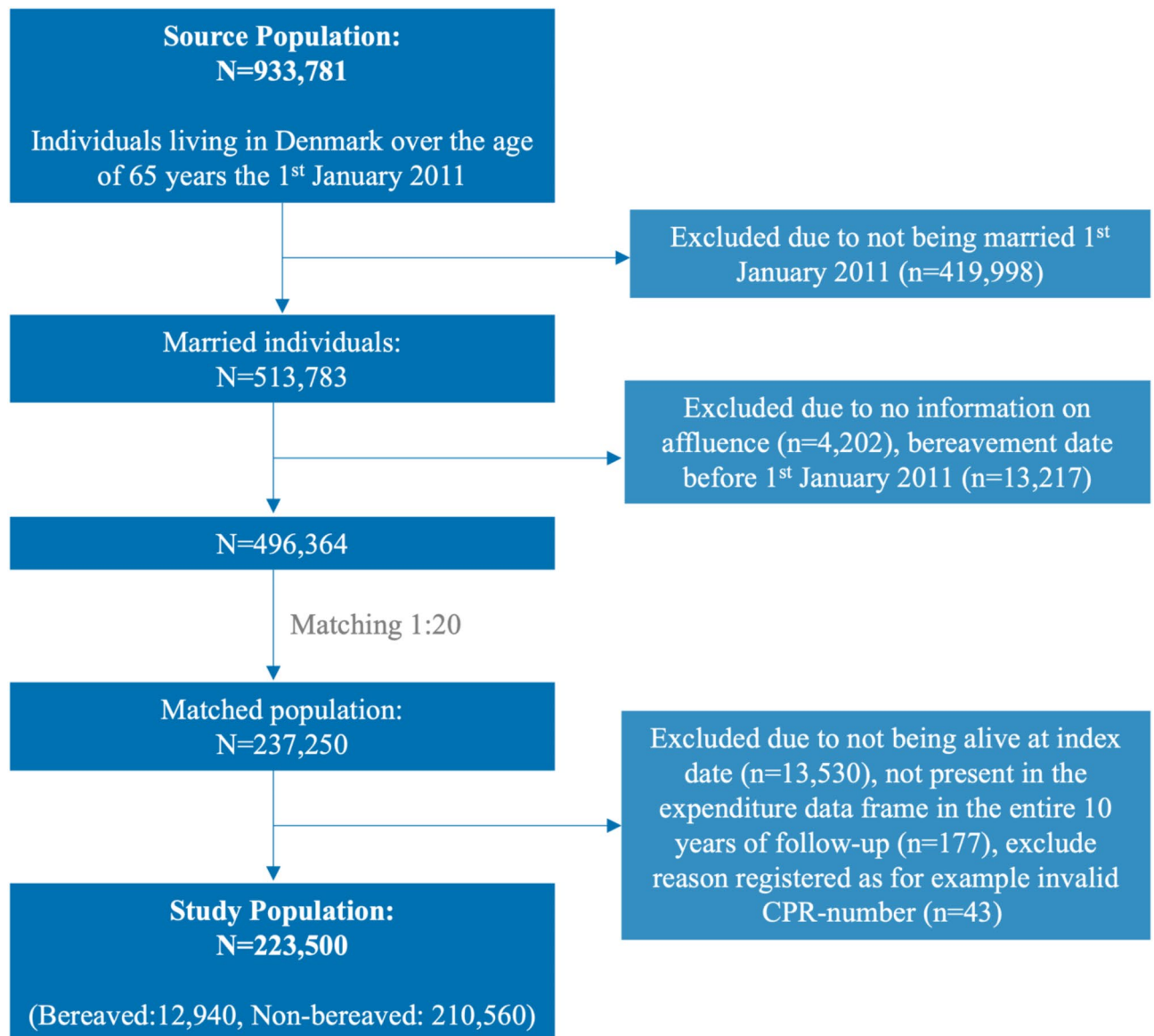


Fig. 1. Flow chart of the study population.

Matching

Matching was used to create comparable groups of bereaved and non-bereaved individuals with shared follow-up periods. Logistic regression propensity score matching was utilized to match each bereaved individual with approximately 20 non-bereaved controls at the date of bereavement (index date) without replacement. The matching criteria included age, sex, number of children, number of comorbidities, immigration status, and affluence index as of January 1, 2011. We excluded controls who did not survive up to the index date, for comparability purposes. Non bereaved individuals were allowed to experience the stressor after the index date. The balance of matching was assessed by evaluating standardized mean differences (SMD), with a threshold of <0.1 indicating balance among covariates. The balance of the matching is presented in Supplementary Fig. 1 (a, b). Mahalanobis distance matching was additionally tested but did not reduce the SMD of the covariates among the two groups.

Statistical methods

The average effect of bereavement

The analysis aimed to estimate the average effect of bereavement on dying from each specific cause for older adults. We performed the analysis using an average treatment effect (ATE) model with a cause specific cox model to account for the competing risk of dying from another cause²⁴. Spousal bereavement was considered the treatment in the model (spousal bereavement $A = 1$, non-bereavement $A = 0$). The average *bereavement* effect was estimated via the *ate* function in R from the *RiskRegression* package, using a G-estimation approach, grounded

in the assessment of counterfactual scenarios (also referred to as a potential outcome model). The ATE for the right censored event time T is given by

$$ATE = \mathbb{E} [\mathbb{E} [Y|A = 1, X] - \mathbb{E} [Y|A = 0, X]]$$

In this study, we let the covariates X be distributed by the sample distribution of the matched population. The estimator for the ATE uses two cox models for competing risks of death (The first model ($j = 1$) predicts the death of interest, and the second model ($j = 2$) predicts death of other causes). The two cox models are given by

$$\lambda_j(t | A, X) = \lambda_{j0}(t) e^{\beta_j X + \gamma_j A}$$

The quantity $\mathbb{E} [Y|A = i, X]$ for $i \in \{0, 1\}$ is estimated by a cox model for competing risk:

$$\hat{\mathbb{E}} [Y|A = i, X] = \int_0^T \hat{\lambda}(s|A = i, X) \hat{S}(s - |A = i, X) ds$$

where $\hat{\lambda}(s|A = i, X) = \hat{\lambda}_{01}(s) e^{\hat{\beta}_1 X + i\hat{\gamma}_1}$ is the cause-specific hazard function of the event happening at time s and $\hat{S}(s - |A = i, X) = e^{-\int_0^s \hat{\lambda}_{10}(t) e^{\hat{\beta}_1 X + i\hat{\gamma}_1} + \hat{\lambda}_{20}(t) e^{\hat{\beta}_2 X + i\hat{\gamma}_2} dt}$ is the event-free survival function.

In this case the average bereavement effect estimates the difference in the average outcome in the case where the whole population is non-bereaved in contrast to the whole population being bereaved. The average bereavement effect is presented in absolute risk differences per 1,000 individuals with corresponding 95% confidence intervals. The confidence intervals were computed using 1,000 bootstraps.

The older adults could either experience 0: not dying (alive), 1: dying from a specific cause, or 2: dying from another cause. The analysis was performed for males and females separately. All analysis were adjusted for the confounders: age, immigration status, affluence index, number of children, number of comorbidities, and healthcare expenditures. In the modelling, we assumed no unmeasured confounding.

Prediction

We restricted the cohort to older adults who had experienced bereavement ($N = 12,940$) for the prediction analysis. We developed prediction models and assessed the prognostic potential of various predictors. We trained the models on 75% of the data and tested on the remaining unseen 25%. For each cause of death, five different prediction models were developed. The first four models were based on a logistic regression model (LRM) with increased number of predictors for each model. The first model was a null model, which assigned the same risk for all individuals in the population based on the mean prevalence of the outcome in the training set (from here on referred to as LRM null). The second model was the simple model with the predictors: age at bereavement, and sex (from here on referred to as LRM simple). The third model included the predictors: age at bereavement, sex, affluence index, number of children, number of comorbidities (from here on referred to as LRM socio+comorbidity). The fourth model included the predictors: age at bereavement, sex, immigration status, affluence index, number of children, number of comorbidities, and each type of healthcare expenditures (prescription, primary care, inpatient and outpatient hospital care, home care, residential care) (from here on referred to as LRM full). This model included restricted cubic splines with three knots to allow for a non-linear relationship between healthcare expenditures and the risk of dying from a specific cause.

We further assessed whether a more flexible machine learning model could increase the predictive capacity of our task compared to the logistic regression model. Specifically, we developed such models using the eXtreme Gradient Boosting (XGBoost) algorithm, which allowed for non-linear effects and interactions between all available predictors. The XGBoost model is a tree-based method that uses decision trees and gradient boosting to create predictions. The XGBoost model included, age at bereavement, sex, immigration status, affluence index, number of children, number of comorbidities, and each healthcare expenditures as predictors (from here on referred to as XGBoost full). For the tuning of hyperparameters for XGBoost we used a 5-fold cross-validation split of the training data. Race ANOVA²⁵ was used as the tuning method to efficiently tune the number of trees, tree depth, number of predictions randomly sampled at each split, minimum number of data points in a node that can be split further, number of iterations that will be performed without improvement and then the learning rate. The combination of hyperparameters that gave the lowest Brier Score was then used for the XGBoost model. We used a maximum entropy design, allowing 30 different candidate values to be tested for each hyperparameter.

The validation of the prediction models was evaluated with the area under the receiver operating characteristic curve (AUC), the Brier Score, the IPA (Index of Predictive Accuracy), calibration plots and potential clinical benefit via decision curve analysis. The AUC quantifies the model's discriminatory ability, measuring the probability of assigning higher risks to individuals with the target outcome compared to those without²⁶.

The Brier Score provides an estimate of the overall prediction error by incorporating both discrimination and calibration. It is defined as the mean squared difference between the predicted probabilities and the actual outcomes. A lower Brier Score indicates superior predictive performance, with a score of 0 representing perfect prediction accuracy, and a score of 1 reflecting complete inaccuracy. The formula for calculating the Brier Score is as follows²⁷:

$$\text{Brier Score} = \frac{1}{N} \sum_{i=1}^N (\text{prediction}_i - \text{outcome}_i)^2$$

The Index of Prediction Accuracy (IPA) measures the overall predictive performance of a model relative to a null model. The null model disregards all predictor variables and assigns the same probability of the outcome to all individuals, corresponding to the prevalence of the outcome in the population. The formula for the IPA is as follows²⁸:

$$\text{IPA} = 1 - \frac{\text{Brier Score(Prediction Model)}}{\text{Brier Score(Null Model)}}$$

The decision curve analysis evaluates the potential clinical utility of a prediction model by quantifying its net benefit across a range of threshold probabilities. This method assesses whether the model improves decision-making compared to default strategies, such as treating all individuals or none, based on varying levels of predicted risk. It integrates the trade-offs between the benefits of true-positive predictions and the harms of false positives - the latter weighted by the odds of the specific threshold probability. More specifically, the formula for the Net Benefit is as follows²⁹:

$$\text{Net Benefit} = \frac{\text{True Positives}}{N} - \frac{\text{False Negatives}}{N} \times \frac{p_t}{1 - p_t}$$

The comparison of the five prediction models based on the mentioned metrics were performed as a stepwise comparison between a more complex model and a simpler one. A delta AUC or Brier Score confidence intervals that did not overlap 0 was considered to reflect a difference in the two rival models being compared. The metrics are presented in percentages (estimates are multiplied by 100). The decision curve analysis was performed to investigate the net benefit at different risk thresholds for each model to gain more insights on clinical utility.

Software

The code is available at GitHub.

Results

Descriptive analysis

Table 1 shows the distribution of baseline characteristics. The study population consisted of a higher proportion of females (61.6%) than males. The age distribution within the study population showed that 20.0% of individuals fell within the 65–69 years range, 27.9% within the 70–74 years range, 24.7% within the 75–79 years range, 16.5% within the 80–84 years range, and 10.9% were aged 85 years or older. Furthermore, the study population was primarily comprised of individuals with Danish origin, accounting for 96.5% of the cohort, while immigrants and their descendants made up the remaining 3.5%. The majority (63.4%) had one or more diseases, with 36.6% without a hospital diagnosis for a disease. The older adults mostly had children, with two children being the most common. The vast majority (76.1%) of the older adults had an average healthcare expenditure of 10–99 DKK/week the year prior to the index date.

The average bereavement effect

Table 2 shows the absolute risks and risk differences for the average bereavement effect on cause-specific mortality within three years for older males and females.

Most of the older adults remained alive during the 3-year follow-up. In total 27,548 older adults died within 3 years of the index date equivalent to 12.3%, with the most common cause of death being cancer and the lowest being psychiatric diseases or suicide. Supplementary Table 2 provides information on the most common causes of death within in cause.

Among older males, the absolute risks of dying from cardiovascular diseases for both older adults who had experienced spousal bereavement (60 deaths/1000 individuals) and older adults who had not experienced bereavement (52 deaths/1000 individuals) showed a risk difference of 8 [95% CI 3;13] additional deaths associated with spousal bereavement per 1000 individuals. For dying from digestive diseases and psychiatric diseases or suicide the results showed smaller absolute risks for both older males who had and had not experienced bereavement with a risk difference of 3 [95% CI 1;5] additional death associated with experiencing bereavement per 1000 individuals. The absolute risk of dying from respiratory diseases were 24 and 28 deaths per 1000 individuals for older males who had not and had experienced bereavement, respectively. Experiencing spousal bereavement was associated with 4 [95% CI 1;8] additional deaths due to respiratory diseases per 1000 individuals compared to older adults who had not experienced bereavement. Experiencing spousal bereavement showed different findings for dying from dementia or Parkinson's, where bereavement was associated with a negative risk difference of 4 less deaths per 1000 individuals [95% CI -6;-2] compared to not experiencing spousal bereavement.

The results for older females showed lower absolute risks. The absolute risks for dying from cardiovascular diseases were less than half of males, with 20 deaths due to cardiovascular diseases for older females who had not experienced bereavement and 25 deaths for older females who had experienced bereavement. Among older females, experiencing bereavement was associated with an additional 5 [95% CI 3;7] deaths due to CVD per 1000 individuals. The absolute risks for dying from psychiatric disease or suicide was 1 death of this cause per 1000 older adults who had not experienced bereavement and 2 deaths for older adults who had experienced bereavement with a risk difference of 1 [95% CI 1;2] deaths associated with bereavement per 1000 individuals.

Variable		Total (<i>n</i> = 223,500)
Bereavement status	Non-bereaved	210,560 (94.2)
	Bereaved	12,940 (5.8)
Sex	Males	85,882 (38.4)
	Females	137,618 (61.6)
Age at bereavement	65–69 years	44,683 (20.0)
	70–74 years	62,253 (27.9)
	75–79 years	55,244 (24.7)
	80–84 years	36,949 (16.5)
	≥ 85 years	24,371 (10.9)
Immigration status	Danish	215,596 (96.5)
	Immigrants or descendants	7,904 (3.5)
Number of comorbidities	No comorbidities	81,853 (36.6)
	1	56,907 (25.5)
	2	36,287 (16.2)
	3	22,018 (9.9)
	≥ 4 comorbidities	26,435 (11.8)
Number of children	No children	25,598 (11.5)
	1	47,675 (21.3)
	2	87,113 (39.0)
	3	44,964 (20.1)
	≥ 4 children	18,150 (8.1)
Affluence index	Lowest	63,531 (28.4)
	Second	55,591 (24.9)
	Third	56,883 (25.5)
	Highest	47,495 (21.3)
Average healthcare expenditures	Less than 10 DKK/week	17,910 (8.0)
	10–99 DKK/week	170,181 (76.1)
	100–499 DKK/week	26,751 (12.0)
	More than 500 DKK/week	8,658 (3.9)

Table 1. Characteristics of the study population in *n*(%) (*N* = 223,500).

Cause of death	Absolute risks [95% CI]		Risk difference [95%CI]
	Non-bereaved	Bereaved	
Males (N= 85,882)			
Cancer (Observed number of deaths = 4,523)	53 [51;54]	54 [49;60]	1 [– 4;7]
Cardiovascular disease (Observed number of deaths = 4,546)	52 [51;54]	60 [55;65]	8 [3;13]
Dementia or Parkinson's disease (observed number of deaths = 1,366)	16 [15;17]	12 [10;15]	– 4 [– 6; – 2]
Diabetes (Observed number of deaths = 374)	4 [4;5]	5 [4;7]	1 [0;3]
Digestive diseases (Observed number of deaths = 447)	5 [5;6]	8 [6;11]	3 [1;5]
Psychiatric diseases or suicide (Observed number of deaths = 121)	1 [1;1]	4 [3;6]	3 [1;5]
Respiratory diseases (Observed number of deaths = 2,093)	24 [23;25]	28 [25;32]	4 [1;8]
Females (N= 137,618)			
Cancer (Observed number of deaths = 3,972)	29 [28;30]	27 [24;29]	– 3 [– 5;0]
Cardiovascular disease (Observed number of deaths = 2,825)	20 [19;21]	25 [23;27]	5 [3;7]
Dementia or Parkinson's disease (Observed number of deaths = 1,197)	9 [8;9]	8 [7;9]	– 1 [– 2;0]
Diabetes (Observed number of deaths = 235)	2 [1;2]	2 [1;3]	0 [0;1]
Digestive diseases (Observed number of deaths = 438)	3 [3;3]	4 [3;5]	1 [0;2]
Psychiatric diseases or suicide (Observed number of deaths = 94)	1 [0;1]	2 [1;3]	1 [1;2]
Respiratory diseases (Observed number of deaths = 1,484)	11 [10;11]	12 [10;14]	1 [0;3]

Table 2. The average bereavement effect on cause-specific death among older adults presented as absolute risks for bereaved and non-bereaved older adults in deaths per 1,000 individuals. Bold risk difference estimates indicate estimates with confidence intervals not overlapping zero. All models were adjusted age, immigration status, affluence index, number of children, number of comorbidities and healthcare expenditures.

The results indicated differences in the risk of dying from cancer, dementia or Parkinson's disease, and diabetes for older adults who had compared to had not experienced bereavement for both males and females. For older females, we additionally found no risk difference for dying from digestive diseases and respiratory diseases.

Predicting cause-specific mortality

We assessed the predictive value of sociodemographic variables and healthcare expenditures for cause-specific mortality within 3 years of spousal bereavement in the restricted data of 12,940 bereaved individuals. Supplementary Fig. 3 shows the AUC, the Brier Scores, and IPAs for all models. The predictive performance varied substantially across different causes of death. The full models showed the highest discriminative ability for dying from diabetes (AUC 88.62% [95% CI 83.31; 93.92%]), dementia or Parkinson's disease (AUC 87.08% [95% CI 82.35; 91.80%]), and respiratory diseases (AUC 81.86% [95% CI 78.03; 85.68%]).

Model comparisons are shown in Fig. 2 which illustrates the differences in the AUC and Brier Score (see Supplementary Table 4 for estimates and 95% CI). The delta AUC shows improved discrimination in all causes of death when comparing the LRM simple to the LRM Null model. Furthermore, adding information on sociodemographic variables and comorbidity did not improve the discrimination apart from dying from CVD (Δ AUC: 2.38 [95% CI 0.26; 4.50]). For dying from CVD adding healthcare expenditures as a predictor also further improved discrimination with a difference in AUC of 2.22 [95% CI 0.65; 3.80] when comparing LRM Full to LRM socio + comorbidities. However, when investigating the prediction accuracy the delta Brier scores showed no indication that adding information on sociodemographic variables and healthcare expenditure improved the predictive performance for all causes of death. For example, when comparing the LRM Full to the LRM socio + comorbidities for dying from cancer the delta Brier score was -0.03 [95% CI -0.12 ; 0.07] indicating no difference in performance.

Machine learning models versus classic models

We compared the XGBoost models against our logistic regression models with all predictors. Across all causes of death, the differences in prediction accuracy were not significant, with all Delta Brier Score confidence intervals including zero. The differences ranged from -0.07 (95% CI: -0.15 to 0.00) for dying from respiratory diseases to 0.05 (95% CI: -0.02 to 0.11) for dying from CVD, indicating that the more complex machine learning approach did not improve upon the logistic regression models with the same predictors. Supplementary Fig. 3 shows the calibration plots for all the models and all causes of death. The calibration plots show very similar performance of the models for all causes of death.

Decision curve analysis

Figure 3 shows the decision curves for the developed models for dying from cancer, CVD, dementia or Parkinson's disease and respiratory diseases. The figure illustrates different scenarios based on the developed models and two additional scenarios *treat all* and *treat none*. *Treat all* describes a scenario, where all individuals will experience treatment without adhering to a prediction model and the *treat none* scenario describes a scenario where all individuals would not experience treatment without adhering to a prediction model. Overall, the figure shows that including the socioeconomic factors and comorbidity as predictors does not improve the net benefit. Including the different healthcare expenditures as predictors improves the net benefit for dying from cancer and dementia or Parkinson's disease. The difference is not as clear for dying from the other causes of death. The XGBoost models does not show higher net benefit apart from dying from dementia or Parkinson's disease between the threshold probabilities around 10–20%. The decision curve analysis further showed that the full models have close to zero net benefit from 10–20% threshold probability and above. Furthermore, it also shows that for dying from diabetes, digestive diseases and psychiatric diseases or suicide the net benefit is zero already from threshold probabilities at around 5%.

Discussion

Main findings

In this extensive cohort study encompassing married older adults, we explored the link between spousal bereavement and cause-specific mortality within a causal framework. We found that older males who had experienced bereavement had higher risks of dying from cardiovascular disease (CVD), digestive diseases, psychiatric disorders or suicide, and respiratory diseases compared to those who did not experience bereavement. Correspondingly, looking at the absolute risk differences for females, we found that experiencing bereavement was associated a higher risk of dying from CVD, psychiatric diseases or suicide and respiratory diseases. The results showed higher absolute risks among bereaved and non-bereaved older males compared to females, but the absolute risks were small for both sexes. This is the first study to our knowledge, that presents absolute risks of dying from a specific cause following spousal bereavement, highlighting the small risk differences for older adults who have not experienced bereavement to those who have.

This study evaluated whether healthcare expenditure data could improve the prediction of cause-specific mortality. While our models showed reasonable discrimination for specific causes of death, such as diabetes, respiratory diseases, and dementia/Parkinson's disease, the addition of healthcare expenditures and other predictors did not improve prediction accuracy compared to simpler models. This suggests that despite having access to comprehensive healthcare utilization data, accurately predicting individual cause-specific mortality remains challenging.

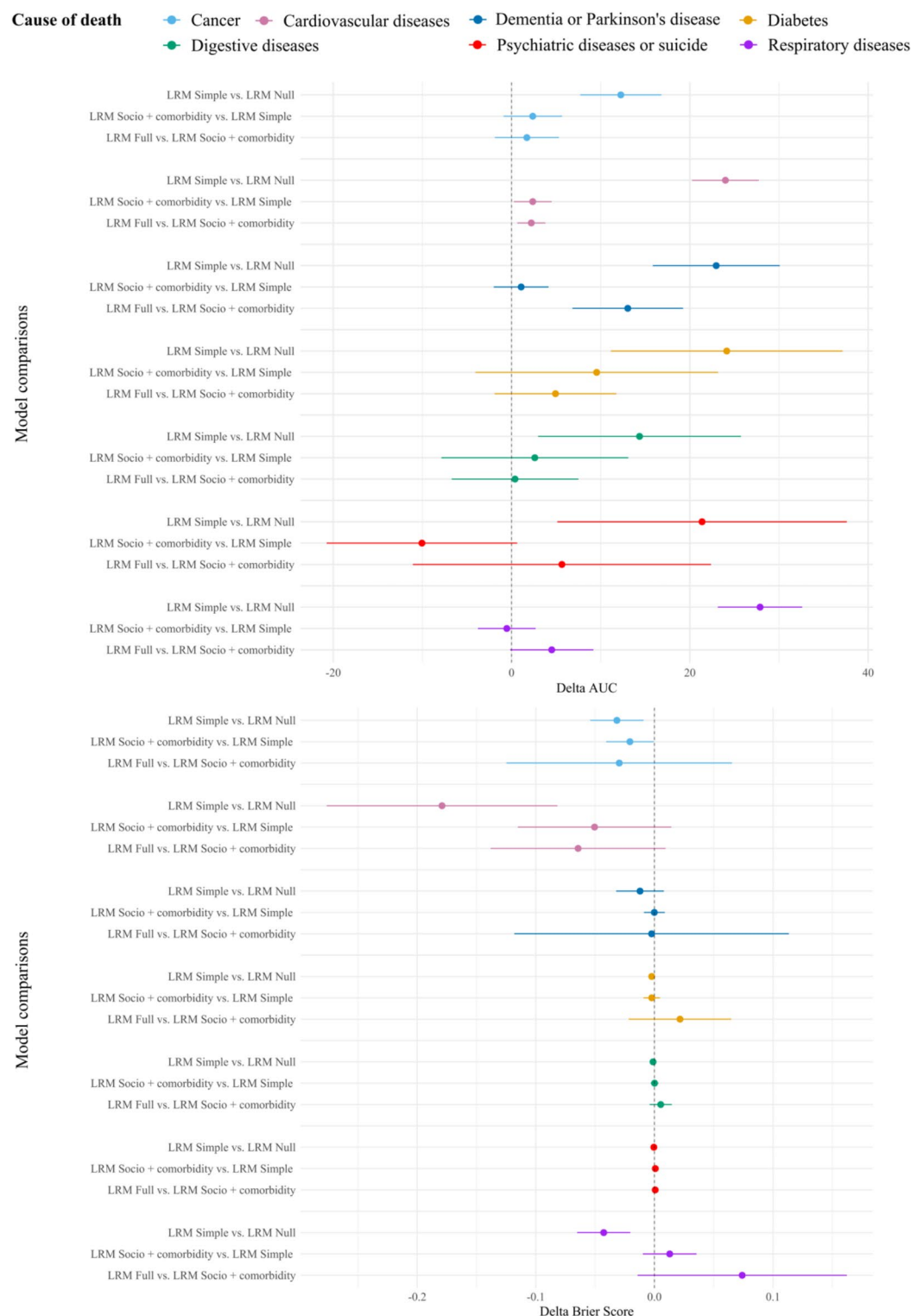


Fig. 2. Model comparison evaluated with delta area under the receiver (AUC) and delta Brier Score with corresponding 95% confidence intervals.

Comparison to previous literature

All previous studies investigating the link between spousal bereavement and mortality have investigated the link between spousal bereavement and cause-specific mortality with regression analysis estimating the magnitude of the association on a relative scale^{13,15–17}. In the present study we have deployed a counterfactual approach using G-estimation in a matched cohort, which supports the hypothesis of causal association. Overall, our findings are in line with the results from previous studies on the association between spousal bereavement and cause-specific

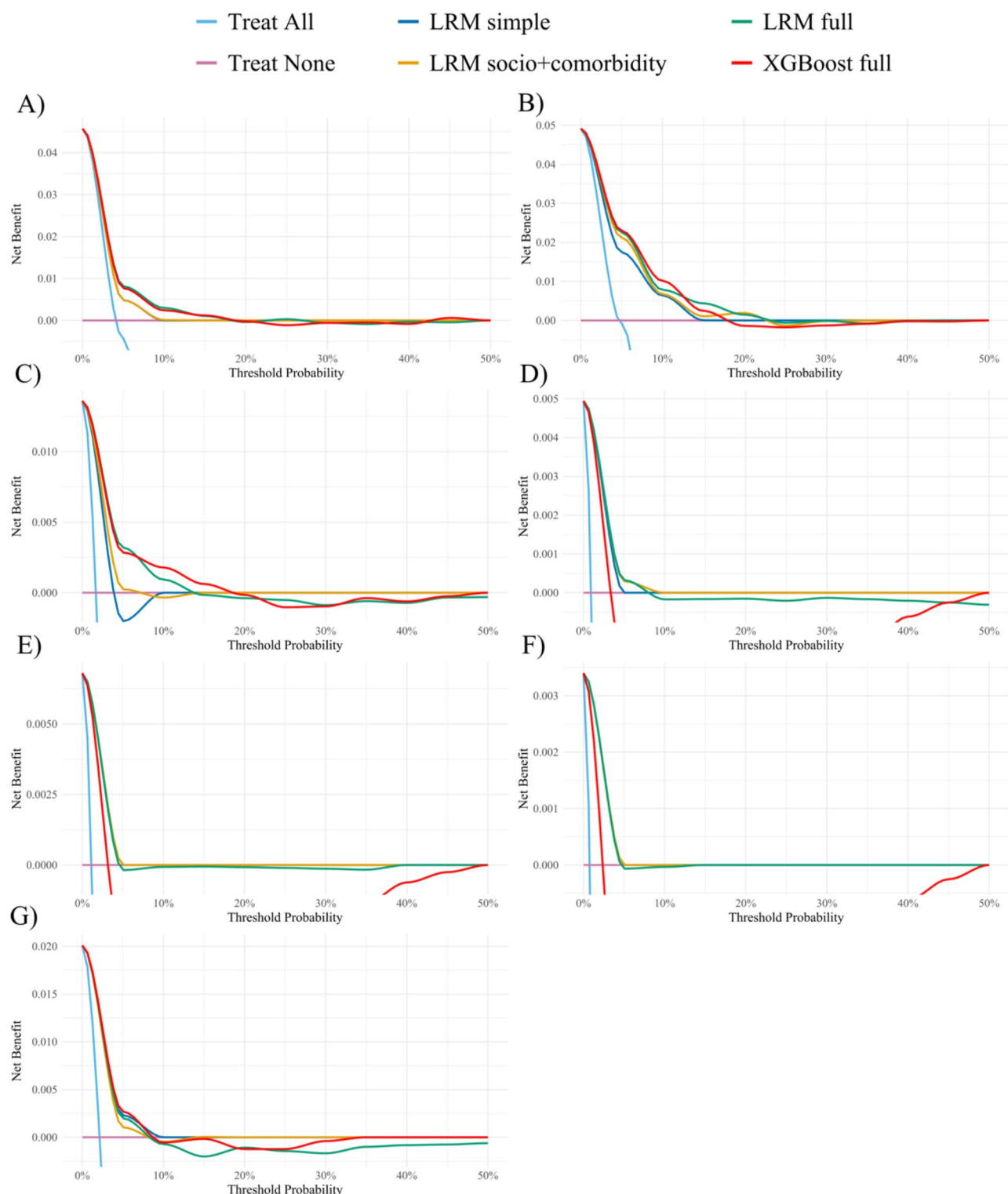


Fig. 3. Decision curve for validation of the prediction models developed to estimate the risk of dying from specific causes based on the logistic regression models (LRM) and extreme gradient boosting (XGBoost) models for dying from (A) cancer, (B) cardiovascular diseases, (C) dementia or Parkinson's disease, (D) diabetes, (E) digestive diseases, (F) psychiatric diseases or suicide, and (G) respiratory diseases. The x-axis shows the probabilities which when exceeded, individuals are classified as high risk of dying from a specific disease within 3 years of the index date. The y-axis shows the proportion of accurately diagnosed and treated individuals after subtracting the weighted false positives for each threshold probability. For example a net benefit of 0.01 means that at this threshold probability, we find one true positive out of 100 individuals without unnecessarily intervening on those.

mortality among older adults^{13,15–17} indicating consistency. Generally the studies adjusted for measurements of sociodemographic and different measures for health status, however one study only adjusted for age and sex¹⁷ and one study lacked information on health status¹⁶. Two Norwegian studies^{16,17} found higher risks of dying from circulatory diseases, respiratory diseases, and external causes for males and females who had experienced bereavement compared to those who had not, but not from cancer. This indicates that the findings of this study also could be applicable to other Scandinavian welfare states. The results for dying from Dementia or Parkinson's disease are in line with previous findings from Elwert and Christakis¹³. Blanner and colleagues did not find an association for neurological diseases and mental disorders among males, which could be explained by the differences in the grouping of diseases as we found a positive association for dying from psychiatric diseases or suicide, but a negative association for dying from Dementia or Parkinson's disease. However, a hypothesized explanation for this finding could be underdiagnosis of Dementia among the bereaved older adults. For dying from cancer, all studies^{13,15–17} found an association, whereas we did not. There are several differences in the study design from the previous studies and ours. We have taken a causal framework approach to the analysis, used matching, and adjusted for several confounders that previous studies have not.

Two previous studies^{10,18}, based on the same source population of older adults who have experienced bereavement as the present study, found that time series of healthcare expenditures were a strong predictor of all-cause mortality. They found evidence that the indicators capture the dynamics of human resilience because they found indication of increased frailty with age. The use of healthcare expenditures was modelled differently in our analysis as we included each type of healthcare expenditure as a separate predictor of an average expenditure one year prior to the index date. The authors of the previous studies used time series of healthcare expenditures and several different measurements of the trends of the expenditures¹⁸ and aggregated time series of healthcare expenditures¹⁰. They found that the healthcare expenditures measured closest to the index date were the most predictive of all-cause mortality. Therefore, limiting the healthcare expenditures to a one-year average in our analysis is a minor possible limitation, but the choice has practical benefits, as the model needs less compute. While the previous studies showed that healthcare expenditures were strong predictors of all-cause mortality their limited performance in our study suggests information on healthcare expenditures may serve better as a general indicator of overall health status rather than capturing the specific patterns needed to predict specific causes of death.

Strengths and limitations

The present study has several strengths. One of the major strengths is the sample size of the study, in which we leveraged nationwide register data with high validity^{20,30,31}.

Another strength is the study design. We chose to exclude unmarried individuals for the probability of everyone experiencing bereavement to be non-zero. This was done with to ensure positivity, which is necessary to draw causal conclusions. The findings are limited to married older adults and are, therefore, not applicable to unmarried cohabitating couples. We assessed the older adults based on the bereavement group they were initially assigned to. Thus, we did not take remarriage into account or that the non-bereaved could experience bereavement during the follow-up. However, remarriage at this age is not common and it would be unlikely to significantly change the direction of the effect.

We used matching of bereaved and non-bereaved as it allowed us to use the average expenditure one year prior to index date for all individuals and to have common follow-ups for all. Furthermore, we used matching to comply with the assumption of exchangeability. The matching has potential to limit the risk of unmeasured confounding, as we hope that the older adults are similar in more ways than in those matched upon. However, there is a risk of not capturing some of the differences between the bereaved and non-bereaved which could bias the results. Furthermore, married couples share endless risk factors such as smoking, diet, sleeping habits that during a lifetime could affect both the dead spouse, but also the bereaved. Some of these risk factors might be reflected in the healthcare expenditures. With respect to the ATE estimation, different ATEs could be obtained under other distributions of the covariates X , which could introduce another source of bias.

The present study has several methodological strengths beyond the matched study population. Using the average treatment effect analysis for estimating the average bereavement effect allows for a natural interpretation of potential outcomes supporting the hypothesis of causal associations. The counterfactual approach reduces the risk of confounding as the effect of bereavement is investigated in the entire study population as what would have been the potential outcome, had all older adults experienced bereavement and what would have happened if they had not experienced bereavement.

No previous study investigating spousal bereavement in older adults has to our knowledge presented their findings as absolute risk estimates. This is a great strength of this study as ratio estimates can be misleading, especially if the risks are small. Absolute risks can be helpful for decision making and for prevention purposes as the potential number of people to help in an intervention is clear.

The healthcare expenditures were used as a proxy for the health of the older adults. In Denmark, where the healthcare system is tax based and out-of-pocket payments only accounts for around 15% of health care expenditures³² (see supplementary text: *Setting*), one could argue that there is free and equal access to the healthcare system and, thus, we would have captured the health of all individuals. However, studies from Denmark have shown that individuals with lower socioeconomic status³³ or from immigrant background³⁴ are less likely to utilize the healthcare system when ill. It is a strength of the present study to include these factors (affluence, immigration status) in the prediction model, as we might be able to capture this complex dynamic.

The present study uses individual person-level data from a cohort of older adults, which is the ideal type of data for developing a prediction model²⁶. We used several metrics to evaluate model performance, whereas several previous studies have been criticised for only evaluating the performance based on the AUC³⁵. In the present study, we assessed both the model's discriminative ability, calibration and presented decision curve

analysis. This is a strength of this study, as poor calibration of a model can make the predictions misleading and thereby reduce the utility of the model³⁶.

Prediction models can have practical applications in either the healthcare system, such as at the general practitioner, or municipality for resource allocation. By identifying individuals at heightened risk of dying for specific causes of death, the model could serve as a tool for prioritizing resources, enabling targeted interventions. However, given our prediction models' limited predictive accuracy, caution about their potential clinical applications is required. The modest performance of even our most complex models suggests that using these predictions for resource allocation or intervention targeting would be premature. Future research should focus on developing more accurate prediction methods before considering clinical implementation.

In conclusion, the findings of our study indicated a higher risk, although still small of magnitude, of death from CVD, digestive diseases, psychiatric diseases or suicide, and respiratory diseases among older males and death from CVD, psychiatric diseases or suicide and respiratory diseases among older females living in Denmark who experienced spousal bereavement, compared to those had not. However, for dying from dementia or Parkinson's disease, the findings showed a lower number of deaths associated with bereavement. The developed prediction models showed low predictive performance for cause-specific mortality even when incorporating comprehensive information on healthcare expenditures and sociodemographic data.

Data availability

The data for the analysis is individual level data from Statistics Denmark. The data can only be shared under specific conditions. According to Danish law, scientific organizations can be authorized to work with data within Statistics Denmark and can provide access to individual scientists inside and outside of Denmark. Data are available via the Research Service Department at Statistics Denmark: www.dst.dk/da/TilSalg/Forskningsservice for researchers who meet the criteria for access to confidential data.

Received: 7 August 2024; Accepted: 14 February 2025

Published online: 20 February 2025

References

- Statistics Denmark. Life expectancy, (2022). <https://www.dst.dk/en/Statistik/emner/borgere/befolkning/middellevetid>. Accessed 20 November 2023.
- Fulop, T. et al. Aging, frailty and age-related diseases. *Biogerontology* **11**, 547–563 (2010).
- Stahl, S. T. & Schulz, R. Changes in Routine Health behaviors following late-life bereavement: a systematic review. *J. Behav. Med.* **37**, 736–755 (2014).
- Kristiansen, C. B. et al. The association of time since spousal loss and depression in widowhood: a systematic review and meta-analysis. *Soc. Psychiatry Psychiatr Epidemiol.* **54**, 781–792 (2019).
- Manzoli, L., Villari, P. & Pirone, M. Marital status and mortality in the elderly: a systematic review and meta-analysis. *Soc. Sci. Med.* **64**, 77–94 (2007).
- Williams, B. R. et al. Intrinsic association of widowhood with mortality in community-dwelling older women and men: findings from a prospective propensity-matched population study. *J. Gerontol. Biol. Sci. Med. Sci.* **66**, 1360–1368 (2011).
- Möller, J. et al. Widowhood and the risk of psychiatric care, psychotropic medication and all-cause mortality: a cohort study of 658,022 elderly people in Sweden. *Aging Ment Health.* **15**, 259–266 (2011).
- Mendes de Leon, C. F., Kasl, S. V. & Jacobs, S. Widowhood and mortality risk in a community sample of the elderly: a prospective study. *J. Clin. Epidemiol.* **46**, 519–527 (1993).
- Katsiferis, A. et al. Sex differences in health care expenditures and mortality after spousal bereavement: a register-based Danish cohort study. *PLoS One.* **18**, e0282892 (2023).
- Katsiferis, A. et al. Dynamical indicators in time series of healthcare expenditures predict mortality risk of older adults following spousal bereavement. *BMC Geriatr.* **22**, 301 (2022).
- Jia, H. & Lubetkin, E. I. Life expectancy and active life expectancy by marital status among older U.S. adults: results from the U.S. Medicare Health Outcome Survey (HOS). *SSM Popul. Health.* **12**, 100642 (2020).
- Jagger, C. & Sutton, C. J. Death after marital bereavement—is the risk increased? *Stat. Med.* **10**, 395–404 (1991).
- Elwert, F. & Christakis, N. A. The effect of widowhood on mortality by the causes of death of both spouses. *Am. J. Public. Health.* **98**, 2092–2098 (2008).
- Chen, H.-K. et al. A prospective cohort study on the effect of sexual activity, libido and widowhood on mortality among the elderly people: 14-year follow-up of 2,453 elderly Taiwanese. *Int. J. Epidemiol.* **36**, 1136–1142 (2007).
- Blanner, C. et al. Widowhood and mortality: a Danish nationwide register-based cohort study. *Epidemiol. Psychiatr. Sci.* **29**, e149 (2020).
- Berntsen, K. N. Trends in total and cause-specific mortality by marital status among elderly Norwegian men and women. *BMC Public. Health.* **11**, 537 (2011).
- Brenn, T. & Ytterstad, E. Increased risk of death immediately after losing a spouse: cause-specific mortality following widowhood in Norway. *Prev. Med.* **89**, 251–256 (2016).
- Katsiferis, A. et al. Machine learning models of healthcare expenditures predicting mortality: a cohort study of spousal bereaved Danish individuals. *PLoS One.* **18**, e0289632 (2023).
- Hernan, M. A. & Robins, J. M. Chapter 1 a definition of causal effect. In: *Causal Inference: What if*. Boca Raton: Chapman & Hall/CRC, 3–4. (2020).
- Helweg-Larsen, K. The Danish Register of causes of Death. *Scand. J. Public. Health.* **39**, 26–29 (2011).
- Lu, T. H. & Using, A. C. M. E. Automatic classification of Medical Entry) software to monitor and improve the quality of cause of death statistics. *J. Epidemiol. Community Health.* **57**, 470–471 (2003).
- Hansen, A. V. et al. Health care expenditure in the last five years of life is driven by morbidity, not age: a national study of spending trajectories in Danish decedents over age 65. *PLoS ONE.* **15**, e0244061 (2020).
- Christensen, B. J., Gøtz, M. & Kallestrup-Lamb, M. Medical spending in Denmark. *Fisc. Stud.* **37**, 461–497 (2016).
- Ozenne, B. M. H. et al. On the estimation of average treatment effects with right-censored time to event outcome and competing risks. *Biometrical J.* **62**, 751–763 (2020).
- Kuhn, M. Efficient grid search via racing with ANOVA models — tune_race_anova. *Fine Tune*, (2014). https://finetune.tidymodels.org/reference/tune_race_anova.html. Accessed 8 January 2024.
- Gerds, T. A. & Kattan, M. W. *Medical risk Prediction Models: With ties to Machine Learning* (CRC Press, Taylor & Francis Group, 2021).

27. Manokhin, V. The Validity and Efficiency of Conformal Prediction. In: *Practical Guide to Applied Conformal Prediction in Python*. United Kingdom, (2023).
28. Gerds, T. A. Index of Prediction Accuracy (IPA), (2022). <https://cran.r-project.org/web/packages/riskRegression/vignettes/IPA.html>. Accessed 1 November 2023.
29. Vickers, A. J. & Elkin, E. B. Decision curve analysis: a Novel Method for evaluating prediction models. *Med. Decis. Mak.* **26**, 565–574 (2006).
30. Thygesen, L. C. et al. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. *Scand. J. Public. Health.* **39**, 12–16 (2011).
31. Pedersen, C. B. The Danish Civil Registration System. *Scand. J. Public. Health.* **39**, 22–25 (2011).
32. OECD. *Extent of Healthcare Coverage*. Paris: OECD.
33. Frølich, A. et al. Multimorbidity, healthcare utilization and socioeconomic status: a register-based study in Denmark. *PLoS ONE.* **14**, e0214183 (2019).
34. Nielsen, M. R. & Jervelund, S. S. Experiences of access to healthcare among newly arrived immigrants in Denmark: examining the role of residence permit. *Scand. J. Public. Health.* 140349482311734. (2023).
35. On behalf of Topic Group ‘Evaluating diagnostic tests and prediction models’ of the STRATOS initiative, Van Calster B, McLernon DJ, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* **17**, 230 (2019).
36. Van Calster, B. & Vickers, A. J. Calibration of risk prediction models: impact on decision-Analytic performance. *Med. Decis. Mak.* **35**, 162–169 (2015).
37. The Danish National Center for Ethics. Registry Research Projects without the Use of Biological Material, (2024). <https://researchethics.dk/information-for-researchers/reporting-to-the-regional-research-ethics-committees/registry-research-projects-without-the-use-of-biological-material>. Accessed 22 December 2024.

Author contributions

All authors have approved the final article. AK and MMBS have full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses. Study concept: AK, SB, MMBS. Study design: AK, SB, MMBS. Acquisition of data: LHM. Mathematical description of analysis: JH. Statistical analysis: MMBS, AK. Interpretation of results: all authors. Drafting of the manuscript: MMBS. Critical revision of the manuscript for important intellectual content: all authors.

Funding

A.K acknowledges the grant Challenge Programme for the project Harnessing the Power of Big Data to Address the Societal Challenge of Aging from Novo Nordisk Foundation (NNF17OC0027812). S.B. acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/X020258/1), funded by the UK Medical Research Council (MRC). This UK funded award is carried out in the frame of the Global Health EDCTP3 Joint Undertaking. S.B. is funded by the National Institute for Health and Care Research (NIHR) Health Protection Research Unit in Modelling and Health Economics, a partnership between UK Health Security Agency, Imperial College London and LSHTM (grant code NIHR200908). Disclaimer: “The views expressed are those of the author(s) and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care.” S.B. acknowledges support from the Novo Nordisk Foundation via The Novo Nordisk Young Investigator Award (NNF20OC0059309) which also supports M.M.B.S., A.K., J.H. and S.B. acknowledges the Danish National Research Foundation (DNRF160) through the chair grant. S.B. acknowledges support from The Eric and Wendy Schmidt Fund For Strategic Innovation via the Schmidt Polymath Award (G-22-63345). Laust Hvas Mortensen is employed at Statistics Denmark, the national Danish Statistics office and is supported in part by grants from the Novo Nordisk Foundation (NNF17OC0027594, NNF17OC0027812) and VILLUM FONDEN (34288). The funding sources had no role in study design, data collection, analysis and interpretation of data, decision to publish or preparation of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Ethics Declaration

All methods were performed in accordance with the relevant guidelines and regulations. Danish legislation allows for register-based research of this type to be conducted without the consent of participants and without ethical committee approval³⁷. The study was conducted according to the rules of the Danish Data Protection Agency. All data was held at Statistics Denmark, which is the Danish national statistical institution.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-90657-1>.

Correspondence and requests for materials should be addressed to M.M.B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025