



Universiteit
Leiden
The Netherlands

Predicting mortality and visualizing health care spending by predicted mortality in Danes over age 65

Hansen, A.V.; Mortensen, L.H.; Ekstrom, C.T.; Trompet, S.; Westendorp, R.

Citation

Hansen, A. V., Mortensen, L. H., Ekstrom, C. T., Trompet, S., & Westendorp, R. (2023). Predicting mortality and visualizing health care spending by predicted mortality in Danes over age 65. *Scientific Reports*, 13(1). doi:10.1038/s41598-023-28102-4

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3764027>

Note: To cite this publication please use the final published version (if applicable).



OPEN Predicting mortality and visualizing health care spending by predicted mortality in Danes over age 65

Anne Vinkel Hansen^{1,2}✉, Laust Hvas Mortensen^{1,3}, Claus Thorn Ekstrøm⁴, Stella Trompet⁵ & Rudi Westendorp^{1,2,3}

Health care expenditure in the last year of life makes up a high proportion of medical spending across the world. This is often framed as waste, but this framing is only meaningful if it is known at the time of treatment who will go on to die. We analyze the distribution of health care spending by predicted mortality for the Danish population over age 65 over the year 2016, with one-year mortality predicted by a machine learning model based on sociodemographics and use of health care services for the two years before entry into follow-up. While a reasonably good model can be built, extremely few individuals have high ex-ante probability of dying, and those with a predicted mortality of more than 50% account for only 2.8% of total health care expenditure. Decedents outspent survivors by a factor of more than ten, but compared to survivors with similar predicted mortality they spent only 2.5 times as much. Our results suggest that while spending in the last year of life is indeed high, this is nearly all spent in situations where there is a reasonable expectation that the patient can survive.

Healthcare expenditure in the last year of life—the “cost of dying”—accounts for between 8.5 and 11.2%¹ of total medical spending. This is even more pronounced at higher ages since this is where mortality is concentrated. A reduction of this expense is an alluring prospect as it would not only reduce the economic burden of an ageing population but potentially also save people useless treatment^{2–5}.

However, the assertion that we spend a lot on those who die is misleading: We can only talk of money wasted on those who go on to die if we can know at the time of starting treatment that the patient will die within a short enough time frame to make treatment frivolous. Recent applications of machine learning to improve prediction of mortality have shown impressive results^{6,7}. Nonetheless, mortality seems to remain to a large extent stochastic^{8,9}—even at very high ages, there is still a substantial probability of living to see another birthday. In fact, the discriminative ability of a number of ‘classic’ health indicators seems to decline with age¹⁰. In a 2018 paper², Einav et al. showed that even a well-performing, state-of-the-art prediction model when predicting one-year mortality for a representative sample of American Medicare enrollees finds very few individuals with more than a 50% risk of dying within a year. It is only with a very high mortality prediction in hand that one could consider to refrain from initiating medical treatment but due to their rarity, individuals with very high mortality risk make up a negligible share of population healthcare expenditure, even though their individual healthcare expenses are high⁴.

In this paper, we attempt to reproduce the analyses of Einav et al.² and explore the distribution of healthcare expenditure by predicted mortality in the Danish population over age 65, using state-of-the-art machine learning methods and the rich national registry data to obtain the best possible prediction of mortality. With this exercise we will explore whether the patterns of mortality and health care expenditure as described for American Medicare enrollees can be found in a different population (covering an entire country) and healthcare system. Next, we explore whether the inclusion of a wider range of socioeconomic variables that are present in the Danish national registries will improve prediction to any material extent, and how the inclusion of the costs of communal care affects the conclusions. The costs of communal care are often ignored when discussing the cost of dying, but these make up a sizeable proportion of the actual societal costs of dying, and the proportion grows with the age of decedents¹¹.

¹Methods and Analysis, Statistics Denmark, , Danmarks Statistik, Sejrøgade 11, 2100 Copenhagen, Denmark. ²Center for Healthy Aging, University of Copenhagen, Copenhagen, Denmark. ³Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. ⁴Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. ⁵Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. ✉email: anviha@dcm.aau.dk

Results

We identified 1,140,242 Danish residents aged 65 and above on Jan 1st 2016 (Fig. 1). After exclusions of 7210 who had immigrated less than two years before start of follow-up and 8971 who were missing from registries, we were left with a study population of 1,124,061 of whom 43,838 (3.9%) died during the year of follow-up. The characteristics of the test population by one-year survival status is shown in Table 1.

Using a machine learning ensemble, the distribution of predicted one-year mortality on the test sample of 374,687 is shown in Fig. 2. While the classification was reasonably good (AUC 0.87), and the distributions of predicted mortalities for survivors and decedents were markedly distinct, only a small proportion (0.6%) had a predicted one-year mortality risk of more than 50%.

Decedents accounted for 13% of 2016 health care expenditure in the test sample (14% of care-related expenditures and 13% of treatment-related expenditures). Figure 3 shows kernel density smoothed mean individual healthcare expenditure per day alive by predicted one-year mortality for decedents and survivors and for total healthcare expenditure, treatment-related expenditure and care-related expenditure. Mean health care expenditure per day alive was higher in decedents than in survivors (1.46 and 0.14 thousand DKK per day respectively). As frail people are more likely to die and also have higher healthcare expenditures, we calculated health care expenditures in a hypothetical population of survivors with the same distribution of predicted mortality as the decedent population had. Mean health care expenditure per day alive in this hypothetical population was 0.57 thousand DKK per day, meaning that decedents outspent survivors by a factor of 10, but outspent equally frail survivors only by a factor of 2.5. Some 39 percent of decedent healthcare expenditure was explained by high spending on those with high ex-ante mortality. The corresponding percentages for treatment- and care-related expenditure, respectively, were 18% and 75%.

Figure 4 shows total healthcare expenditure by predicted mortality. While spending was clearly higher at higher predicted mortalities, the group with a predicted mortality of more than 50% only accounted for 2.8% of total spending. The numbers for treatment and care were 1.8% and 4.4% respectively. While 75% of total treatment-related spending occurred at a predicted mortality of less than eight percent, care-related expenditure

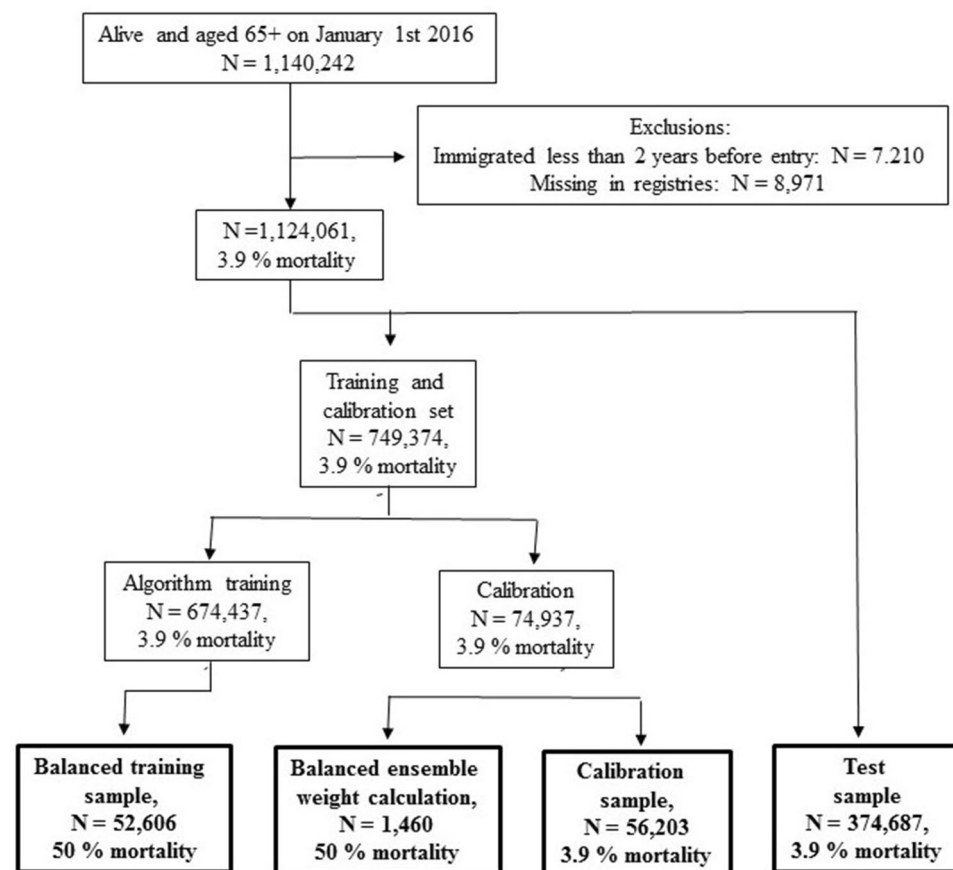


Figure 1. Chart describing inclusion in the study population, and data flow into the prediction algorithm. Individuals were included in the national population study if they were alive and aged 65 + on January 1st 2016 and had been living in Denmark for two years prior to this baseline date. A random sample of two thirds was selected, and a prediction model was trained on this sample, predicting individual-level mortality risk in the year after baseline based on characteristics observed during the two years before baseline. The prediction model was then used to predict one-year mortality risk for the remaining test sample, and this sample was followed up for mortality and healthcare expenditure for up to one year after baseline.

	Survivor		Decedent	
	360,074		14,613	
Age (n, %)				
64–69	131,478	37	1783	12
70–74	94,516	26	2104	14
75–79	61,630	17	2340	16
80–84	38,731	11	2663	18
85–89	22,210	6	2713	19
90–94	9129	3	2075	14
95–99	2143	1	786	5
100 +	237	0	149	1
Sex (n, %)				
Male	164,695	46	7018	48
Female	195,379	54	7595	52
Danish (n, %)				
Danish	343,634	95	14,043	96
Immigrant	16,440	5	570	4
Education (n, %)				
Elementary shool	132,303	37	6950	48
High school, vocational training	141,269	39	4570	31
Further education	78,135	22	1945	13
Unspecified	8367	2	1148	8
Household type (n, %)				
Cohabiting/Married	218,538	61	4864	33
Divorced	40,764	11	1865	13
Single	25,564	7	1817	12
Widow/er	75,208	21	6067	42
Multimorbidity (n, %)				
0	50,591	14	397	3
1–2	120,201	33	2,130	15
3–4	92,664	26	3,666	25
5 +	96,618	27	8,420	58
Received care services				
No	335,932	93	8638	59
(n, %)				
Home care only	18,130	5	3693	25
Nursing home	6012	2	2282	16
Predicted 1-year mortality (mean, IQR)	0.03 (0.01–0.03)		0.21 (0.06–0.34)	

Table 1. Characteristics of the test sample used for calculating health care expenditures by predicted mortality.

was concentrated among those with moderately high predicted mortality, with 75% of expenditure concentrated at predicted mortalities over 11 percent.

Discussion

Decedents accounted for 13% of yearly healthcare expenditure at age 65 and above, but only 2.8% was spent on those who, according to our machine learning model, had a likelihood of dying of more than 50%. While the mean healthcare expenditure per day alive on a decedent was ten times that of a survivor, when comparing to an equally frail population of survivors, the mean expenditure per day alive on a decedent was only 2.5 times higher. The main strength of the study is the availability of data for an entire population, with rich health care and sociodemographic predictor data and registry coverage of 97% of all healthcare expenditure¹², as well as the inclusion of communal care in addition to treatment. As healthcare expenditure in Denmark is tax-funded, differences will not be artefacts of differential insurance coverage and rates. Individual-level expenditure data, however, may be misestimated to some extent: Hospital costs are DRG rates which are averages and may not entirely correspond to the actual cost of treatment, and the computation of individual-level expenditure on nursing home and home care involve some amount of estimation and imputation. The study deals only with expected mortality at baseline, which may arguably be a limited indicator of cost-efficiency of healthcare spending, and other measurements such as quality-adjusted life years could have been taken into account.

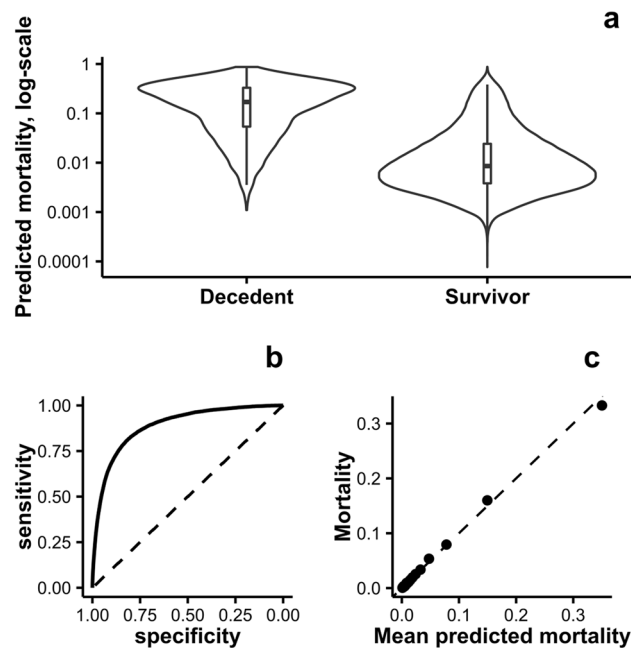


Figure 2. Description of the one-year mortality predictions for the test sample. Violin plot of the distribution of predicted mortalities for survivors and decedents (a), ROC curve (b), calibration plot (c).

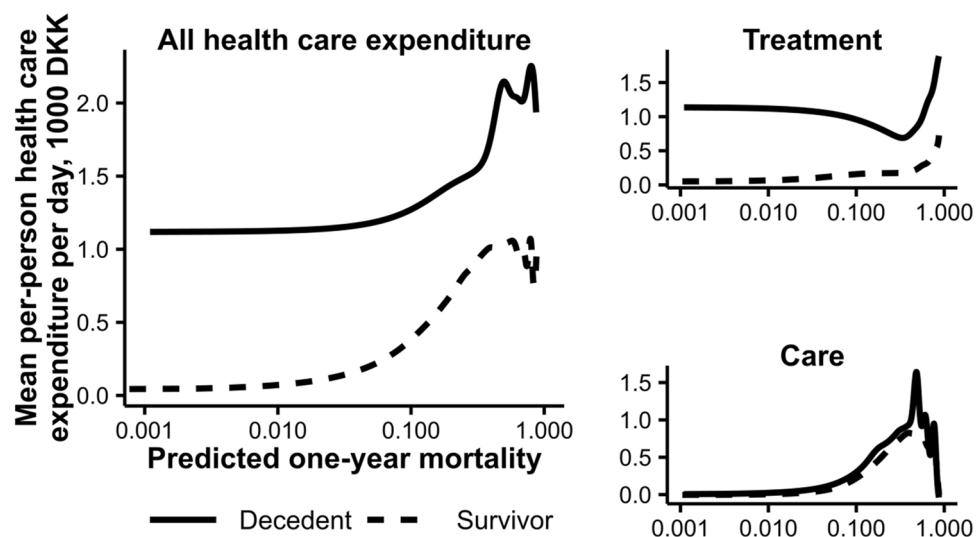


Figure 3. Kernel smoothed per-person healthcare expenditure per day by predicted one-year mortality and type of expenditure.

The distribution of predicted mortalities resemble that estimated² for American Medicare enrollees. The inclusion of a wider array of personal characteristics has not improved prediction materially, as our AUC is essentially the same as that of the Medicare study—a result that compares reasonably well to what other studies have achieved^{16,17,18}, particularly considering the relatively wide time horizon of prediction for our study. The very low proportion with high predicted mortality might be due to essential randomness in mortality, the accrual of health-impacting events after start of follow-up, or due to shortcomings of the data available. But while we absolutely might point to health indicators that were not available for the study there are indications^{10,17,18} that these may not improve mortality prediction that much.

The mass of treatment costs is concentrated at low predicted mortalities in a pattern resembling that of Einav et al.² Care-related costs, conversely, are concentrated at higher mortalities and increase more markedly with increasing mortality, whereas the costs of treatment among decedents actually decrease up to a predicted mortality of about 30%. This is not surprising—predicted mortality is a proxy for frailty and thus for the need

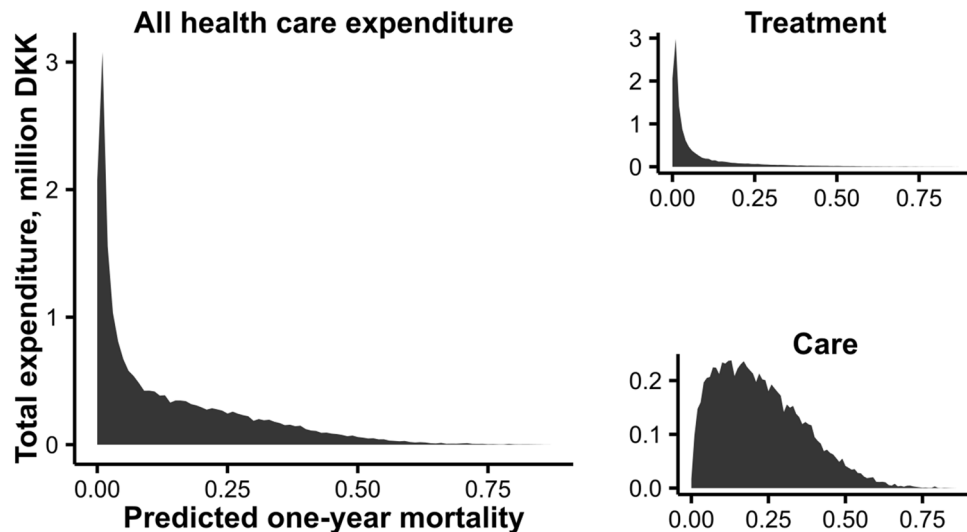


Figure 4. Healthcare expenditure by predicted one-year mortality expenditure and type of expenditure.

for communal care, and the need for care is likely to change less as the result of health-impacting events over the course of follow-up. It is interesting that we see a decline with predicted mortality in treatment-related expenditure per day alive for decedents. This was not observed for the American population and may reflect different medical culture in Denmark and the US, but the different prediction algorithms might also be part of the explanation—treatment-related expenditure decreases with age in Danish decedents¹¹, and if a high predicted mortality is more reflective of age and frailty in our algorithm than for the American data, that might explain the difference.

At similar predicted mortalities, there is little difference between the care-related expenditure per day alive of decedents and survivors. The treatment-related expenditure of decedents is much higher than that of survivors, although the differences are lower at higher mortalities. This pattern may in part be explained by the passage of time—by the time a person dies, their health has likely deteriorated since their status at entry, and it seems likely that a person who dies at low predicted mortality will have experienced some dramatic health event requiring treatment, while death at higher predicted mortality might be a more direct continuation of patterns already established by the time of entry. Also, a person with low predicted mortality might be a better candidate for treatment, being less frail. But to the extent that the difference between survivors and decedents at the same mortality is not due to curveball events, it might be seen as the “true” cost of dying.

Thus, nearly all healthcare expenditure occurs in situations where there is a reasonable expectation that the patient can survive, and so the concept of “the cost of dying” is confounded by frailty: We spend more on the frail, and the frail are more likely to die—but not certain to do so, at least within a relevant time frame. This underlying frailty, operationalized as high predicted one-year mortality accounted for 39% healthcare expenditure in the last year of life in Denmark, an estimate in line with that in American Medicare enrollees². The idea of a potential for reductions in health care expenditure at the end of life is enticing, and it seems possible to find groups that could benefit from a switch to a palliative course of treatment. Still, our results, along with those of our model paper, add to a list of arguments for why it might be illusory to reduce healthcare expenditure much by cutting the cost of dying. The proportion of spending occurring at the end of life is lower than has previously been reported¹, decedents make up a relatively small share of high-cost individuals³, rising levels of demand drive increasing health care costs in ageing populations at least as much as the cost of dying¹⁹, and high end-of-life costs seem driven more by multimorbidity than last-ditch lifesaving efforts^{1,11,20}. Our study design does not touch upon the question of individual treatment effects—whether specific treatments improve survival for specific individuals—and it may be that better methods than ours can detect high-mortality subgroups, but it seems unlikely for such subgroups to be large enough that costs reductions there could matter on the scale of a national budget.

Methods

The use of personal data in this study followed Danish data protection legislation. Treatment of the data at Statistics Denmark is legal according to the Act on Statistics Denmark²¹ and the General Data Protection Regulation (GDPR) art 6 s 1 ss e²². Statistics Denmark has the right to process individual-level registry data. Per GDPR art 14 s 5 ss b²², obtaining individual consent for the use of registry data in research is not required.

The study design is illustrated in Fig. 1. The study population consisted of all individuals of age 65 and above who had been living in Denmark for two years at the point of entry into the study (January 1st 2016). We collected personal characteristics from individually linkable national registries for the last two years before entry and followed up for death or emigration for one year after entry, as well as for use of healthcare services in the follow-up period (Supplementary information 1).

A machine-learning algorithm to predict one-year survival was trained on a training sample consisting of two thirds of the population. The algorithm was an ensemble of three methods: Boosting, Random Forest and Lasso.

Sub-samples of 2.5% and 7.5% of the training sample were set aside for ensemble weight calculation and calibration, the remaining training sample was reduced to a balanced sample with a 50–50 distribution of decedents and survivors by random selection of a subsample of survivors, and the three learners were trained on the balanced training sample, with parameter tuning by five-fold cross-validation. Having trained the individual learners, we created an ensemble by weighting them together, using the ensemble weight calculation sample (after reducing this to a balanced sample) to compute weights by computing predicted values for the three learners and fitting a linear combination of these to the observed values. As the prediction ensemble was trained on balanced data, we re-balanced it, using Bayes' rule. Following our model paper, we would have used the calibration sample to fit a third-degree polynomial in the predicted values to observed mortality, but as we observed better calibration simply using Bayes' rule, we abandoned that approach. The prediction model was trained using R software and the tidymodels framework²³, and the ranger²⁴, xgboost^{25,26} and glmnet²⁷ packages.

The personal characteristics were sex, age, country of origin, education, marital status and household size, municipality, variables on income and financial assets, use of primary healthcare grouped by specialization, number of hospital contacts by ICD-10 chapter of the main diagnosis for the contact, number of prescriptions by five-digit ATC code, amount of home care (personal and practical help) being provided by the municipality and an indicator for admission to a nursing home. The financial variables were collected on a yearly basis for the last two years before entry, all other time-varying variables were collected on a quarterly basis for the last two years before entry. In order to reduce the size of the predictor space, we performed principal component analysis on the quarterly-level prescription data and reduced to the first 66 components.

The trained algorithm was used to predict one-year survival for the test sample of one third of the population, and the distribution of healthcare expenditure over the year of follow-up by predicted mortality at one year was examined. Healthcare expenditure was defined either treatment-related: The costs of hospitals, primary care and prescription drugs; or care-related: The costs of home care and residential care. Healthcare and eldercare in Denmark is primarily taxpayer-funded, so the information available in registries accounts for 97% of healthcare expenditure in Denmark¹². Individual-level expenditures for the period of follow-up were computed, as well as individual-level expenditure per day of follow-up, for total health expenditure and for treatment- and care-related expenditures. The methods are described in¹¹.

In order to assess the proportion of health care expenditure explained by predicted mortality, we reweighted the survivor population to have a similar distribution of predicted mortality as the decedent population and compared mean spending per day alive in the weighted survivor population to mean spending among decedents. The weights were computed, using kernel density estimates of the densities among decedents and survivors with a survivor with an estimated survival probability of p being given a weight of $k_d(p)/k_s(p)$ where k_d is the density function estimated for decedents and k_s is the density function estimated for survivors.

Data availability statement

Due to restrictions in Danish law, the confidential health care data used in this study can only be accessed through Statistics Denmark, the state organization holding the rights to the data. Danish 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. The authors of this study had no special access privileges others would not have.

Received: 11 October 2022; Accepted: 12 January 2023

Published online: 21 January 2023

References

- French, E. B. *et al.* End-of-life medical spending in last twelve months of life is lower than previously reported. *Health Aff. (Millwood)* **36**, 1211–1217 (2017).
- Einav, L., Finkelstein, A., Mullainathan, S. & Obermeyer, Z. Predictive modeling of U.S. health care spending in late life. *Science* **360**, 1462–1465 (2018).
- Aldridge, M. D. & Kelley, A. S. The myth regarding the high cost of end-of-life care. *Am. J. Public Health* **105**, 2411–2415 (2015).
- Zhang, B. *et al.* Health care costs in the last week of life: Associations with end-of-life Conversations. *Arch. Intern. Med.* **169**, 480–488 (2009).
- Smith, S., Brick, A., O'Hara, S., Normand, C. Evidence on the cost and cost-effectiveness of palliative care: A literature review. at <<https://journals.sagepub.com/doi/full/https://doi.org/10.1177/0269216313493466>> (2014).
- Rajkomar, A. *et al.* Scalable and accurate deep learning with electronic health records. *NPJ Digit. Med.* **1**, 1–10 (2018).
- Zhang, G., Xu, J., Yu, M., Yuan, J. & Chen, F. A machine learning approach for mortality prediction only using non-invasive parameters. *Med. Biol. Eng. Comput.* <https://doi.org/10.1007/s11517-020-02174-0> (2020).
- Coggon, D. & Martyn, C. Time and chance: The stochastic nature of disease causation. *Lancet* **365**, 1434–1437 (2005).
- Smith, G. D. Epidemiology, epigenetics and the 'Gloomy Prospect': Embracing randomness in population health research and practice. *Int. J. Epidemiol.* **40**, 537–562 (2011).
- Kusumastuti, S., Gerdts, T. A., Lund, R., Mortensen, E. L. & Westendorp, R. G. J. Discrimination ability of comorbidity, frailty, and subjective health to predict mortality in community-dwelling older people: Population based prospective cohort study. *Eur. J. Intern. Med.* **42**, 29–38 (2017).
- Hansen, A. V., Mortensen, L. H., Trompet, S. & Westendorp, R. 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: Medical spending in Denmark. *Fisc. Stud.* **37**, 461–497 (2016).
- Tabak, Y. P., Sun, X., Nunez, C. M. & Johannes, R. S. Using electronic health record data to develop inpatient mortality predictive model: Acute laboratory risk of mortality score (ALaRMS). *J. Am. Med. Inform. Assoc.* **21**, 455–463 (2014).

14. Makar, M., Ghassemi, M., Cutler, D. M. & Obermeyer, Z. Short-term mortality prediction for elderly patients using medicare claims data. *Int. J. Mach. Learn. Comput.* **5**, 192–197 (2015).
15. Pirracchio, R. *et al.* Mortality prediction in intensive care units with the super ICU learner algorithm (SICULA): A population-based study. *Lancet Respir. Med.* **3**, 42–52 (2015).
16. Awad, A., Bader-El-Den, M., McNicholas, J. & Briggs, J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int. J. Med. Inf.* **108**, 185–195 (2017).
17. Kusumastuti, S., Rozing, M. P., Lund, R., Mortensen, E. L. & Westendorp, R. G. J. The added value of health indicators to mortality predictions in old age: A systematic review. *Eur. J. Intern. Med.* **57**, 7–18 (2018).
18. Zeltzer, D. *et al.* Prediction accuracy with electronic medical records versus administrative claims. *Med. Care* **57**, 551–559 (2019).
19. Kjellberg, J. Fremtidens sundhedsudgifter: En analyse af sund aldring, ”steeping” og teorien om alder som ”red herring”. (2020).
20. Davis, M. A., Nallamothu, B. K., Banerjee, M. & Bynum, J. P. W. Identification of four unique spending patterns among older adults in the last year of life challenges standard assumptions. *Health Aff. (Millwood)* **35**, 1316–1323 (2016).
21. Indenrigs- og Boligministeriet. *Bekendtgørelse af lov om Danmarks Statistik. LBK nr 610 af 30/05/2018*, (2018).
22. Europa-Parlamentets og Rådets forordning (EU) 2016/679 af 27. april 2016 om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger og om fri udveksling af sådanne oplysninger og om ophævelse af direktiv 95/46/EF (generel forordning om databeskyttelse) (EØS-relevant tekst). *OJ L* **119**, (2016).
23. Kuhn, M. & Silge, J. *Tidy Modeling with R*. (2020).
24. Wright, M. N. & Ziegler, A. **ranger**: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**, (2017).
25. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 <https://doi.org/10.1145/2939672.2939785> (2016).
26. Chen, T. & He, T. xgboost: eXtreme Gradient Boosting. 4.
27. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

Acknowledgements

This study was funded by an unconditional research grant from Novo Nordisk Fonden (NNF17OC0027812).

Author contributions

Study conception and design: A.I.H., L.H.M., R.W., S.T. Methodology: A.I.H., C.T.E. Data collection and analysis: A.I.H. Writing, discussion, revision: A.I.H., C.T.E., L.H.M., R.W., S.T. All authors have read and agreed to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-28102-4>.

Correspondence and requests for materials should be addressed to A.V.H.

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) 2023