OPEN

Epidemiology Publish Ahead of Print

DOI:10.1097/EDE.0000000000000920

Dynamic prediction of survival in cystic fibrosis: A landmarking analysis using UK patient registry data

Ruth H Keogh[1], Shaun R Seaman[2], Jessica K Barrett[2], David Taylor-Robinson[3], Rhonda Szczesniak[4].

[1]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London, United Kingdom.

[2]Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom.

[3]Department of Public Health and Policy, Farr Institute@HERC, University of Liverpool, Liverpool, United Kingdom.

[4]Division of Biostatistics and Epidemiology and Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center; Department of Paediatrics, University of Cincinnati, Cincinnati, Ohio, United States.

Corresponding author: Ruth Keogh., Mailing address: Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London, United Kingdom.Email: ruth.keogh@lshtm.ac.uk, Phone: +44 20 7927 2570

Running head: Dynamic prediction of survival in cystic fibrosis

Conflicts of interest: There are no conflicts of interest.

Sources of financial support: RHK and DTR were supported by a Medical Research Council Fellowships (MR/M014827/1 (RHK), MR/P008577/1 (DTR). SRS was supported by MRC Grant MC_UU_00002/10. JKB was supported by a Medical Research Council unit programme (MC_UU_00002/5). RS was supported by US NIH Grant K25 HL125954.

ACCEPTED

Downloaded from https://journals.lww.com/epidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3z484KGaITVFBAZIS+s7Y29myPseAR9RdGwhXLlVgwE=gDRTHMsxw== on 11/16/2018

Data and code: This work used anonymised data from the UK Cystic Fibrosis Registry. Data are available following application to the Registry Research Committee. https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/apply-for-data-from-the-uk-cf-registry. Example code for obtaining estimated survival probabilities from the final model presented is provided at https://github.com/ruthkeogh/landmark_CF. Code used in the analyses is also provided at the same webpage. Further details are given in the Supplementary Materials.

Acknowledgements: We thank people with cystic fibrosis and their families for consenting to their data being held in the UK CF Registry, and NHS teams in CF centres and clinics for the input of data into the Registry. We also thank the UK Cystic Fibrosis Trust and the Registry Steering Committee for access to anonymized UK CF Registry data. The analyses presented in this paper use a data set resulting from a cleaning process undertaken by the CF Epidemiological Network, which was funded by a Strategic Research Centre Grant from the Cystic Fibrosis Trust.

Keywords: Cox regression, Cystic fibrosis, Dynamic prediction, Landmarking, Longitudinal data, Patient registry, Personalised prediction, Survival.

Copyright © 2018 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ACCEPTED

ABSTRACT

Background

Cystic fibrosis (CF) is an inherited, chronic, progressive condition affecting around 10,000 individuals in the UK and over 70,000 worldwide. Survival in CF has improved considerably over recent decades and it is important to provide up-to-date information on patient prognosis.

Methods

The UK Cystic Fibrosis Registry is a secure centralized database, which collects annual data on almost all CF patients in the UK. Data from 43,592 annual records from 2005-2015 on 6181 individuals were used to develop a dynamic survival prediction model that provides personalized estimates of survival probabilities given a patient's current health status using 16 predictors. We developed the model using the landmarking approach, giving predicted survival curves up to 10 years from ages 18 to 50. We compared several models using cross-validation.

Results

The final model has good discrimination (C-indexes 0.873, 0.843, 0.804 for 2-, 5-, 10-year survival prediction) and low prediction error (Brier scores 0.036, 0.076, 0.133). It identifies individuals at low and high risk of short- and long-term mortality based on their current status. For patients aged 20 during 2013-2015, for example, over 80% had a greater than 95% probability of 2-year survival and 40% were predicted to survive 10 years or more.

Conclusions

Dynamic personalized prediction models can guide treatment decisions and provide personalized information for patients. Our application illustrates the utility of the landmarking approach for making the best use of longitudinal and survival data and shows how models can be defined and compared in terms of predictive performance.

ACCEPTED

INTRODUCTION

Cystic fibrosis (CF) is an inherited, chronic, progressive condition affecting around 10,000 individuals in the UK and over 70,000 worldwide.[1,2] In the United Kingdom (UK) CF affects about 1 in 2500 live births [3]. Children with CF are generally diagnosed in the first few months of life, with universal newborn screening implemented in 2007 in the UK, though some people with milder phenotypes are diagnosed into adulthood.[4]

Survival in CF has improved considerably over recent decades. Of individuals born around 1970, over half died before reaching their mid- to late teens.[5,6] By contrast, the estimated median survival age for a person born with CF today in the UK is 48 for males and 44 for females.[1,7] It is important to be able to provide patients with up-to-date information on their prognosis, and to provide clinicians with information to guide treatment decisions, including listing for lung transplantation.

Data from national CF patient registries with longitudinal measures of health status and long term follow-up have created the opportunity to develop models for predicting survival based on individual characteristics.[8,9] Although there have been many studies of factors associated with survival in CF (see Buzetti et al.[10] and MacNeill[3] for overviews), fewer have focused on prediction. We identified three models for survival prediction in UK patients, but all are based on small samples or subsets of patients.[11–13] Survival prediction models in CF have been developed using national patient registries by Liou et al.[14] and Mayer-Hamblett et al.[15] (United States), Aaron et al.[16] (Canada), and Nkam et al.[17] (France). Until recently there have been (to our knowledge) no detailed studies of survival using the UK CF Registry. Keogh et al.[18] provided estimates of survival using UK CF Registry data given the baseline characteristics of sex, genotype, and age of diagnosis. In this paper we develop a model for personalized prediction of survival in the UK making use of time-dependent measures of health status.

ACCEPTED

The aims of this article are twofold. Our first aim was to use data from the UK CF Registry to develop a dynamic survival prediction model that provides estimates of the probability of short-term, mid-term and long-term survival given a patient's current and past health status.[19] We used the landmarking approach applied to UK CF Registry data on adults from 2005-2015, [20,21] giving predicted survival curves up to 10 years from each landmark age, which can be any age post-diagnosis. The model therefore provides predictions for individuals living with the CF who already survived to a given age. The model is dynamic in that it enables predictions to be updated over time, using updated measures of time-dependent predictors alongside a patient's current age. Our second aim was to provide an example for other researchers of how to develop a dynamic prediction model using landmarking, illustrating the utility of this approach for making the best use of longitudinal and survival data, and showing how different models can be defined and compared in terms of their predictive performance.

METHODS

Design and data source

We undertook a landmarking analysis using data from the UK CF Registry, a national, secure database sponsored and managed by the Cystic Fibrosis Trust.[19] The Registry was established in 1995 and records demographic data and longitudinal health data on nearly all people with CF in the UK, to date capturing data on over 12,000 individuals. NHS Research Ethics approval has been granted for the collection of data into the Registry. Each patient or their parent provided written informed consent for collection of data in the Registry and use of pseudonymized data in research. In the UK, CF patients are treated in specialist centres and data for the Registry are collected in a standardized way at designated (approximately) annual visits. Data collected cover over 250 variables in several domains, alongside mortality data. We restricted our analyses to a set of 17 variables (Table 1) recorded routinely in the Registry and previously found to be associated with survival, based on a review of the literature.

ACCEPTED

[3,10,11,13,15–17,22–28] This set consists of three baseline variables – sex, genotype (F508del alleles), and age of diagnosis—as well as calendar year, and 13 internal time-dependent variables: forced expiratory volume in 1 second as percentage predicted (FEV1%); forced ventricular capacity as percentage predicted (FVC%); height; weight; infection status for four organisms (*Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Burkholderia cepacia*, Methicillin-resistant *Staphylococcus aureus* [MRSA]); CF-related diabetes; pancreatic insufficiency; days in hospital on intravenous (IV) antibiotics; days at home on IV antibiotics; and other hospitalization. We calculated FEV1% and FVC% using the Global Lung Initiative (GLI) equations.[29] We investigated using BMI instead of weight and height, but found that models including weight and height separately were better fitting, based on Akaike's Information Criterion.[30] The two variables for days on IV antibiotics are used as surrogate indicators for pulmonary exacerbations.[31,32]

Analyses are based on follow-up during the study period 2005-2015, so that some individuals have at least 10 years of follow-up, enabling estimation of survival up to 10 years. We therefore excluded individuals who died or were lost to follow-up before 2005. In order to focus on adults, we only used data on individuals from age 18 onwards during the study period.

The landmarking approach

The landmarking approach for dynamic prediction of survival was first described by van Houwelingen.[20] A detailed account is provided by van Houwelingen and Putter.[21] In brief, at a given age (a 'landmark age') from which a prediction is to be made, the data are restricted to individuals who have not yet had the event (in this case, death) or been censored. Values of predictor variables available up to the landmark age are used as covariates in a model for the probability of survival up to some time horizon, conditional on survival to the landmark age. Typically, the focus is on survival to a single time horizon (t_{hor}), e.g. 2 years after the

ACCEPTED

landmark age ($t_{hor} = 2$), and censoring is imposed at t_{hor} so that only events up to that time are used in the survival analysis. For a chronic condition like CF, however, it is of interest to study survival to several time horizons. We use the Cox model and its extensions to model survivor curves up to 10 years after each landmark age.

Landmark data sets were created from landmark ages $l = 18, \ldots, 50$ (eFigure 1; http://links.lww.com/EDE/B407, eTable 1; http://links.lww.com/EDE/B407, eAppendix 1; http://links.lww.com/EDE/B407). Data on individuals aged over 50 are sparse. The lth landmark data set included all individuals known to be alive at age l during 2005-2015, who had not received a transplant prior to age l, who were diagnosed with CF before age l, and who joined the Registry before age l. Individuals lost to follow up before age l were excluded. We excluded people who received a transplant prior to age l because the variables of importance for survival in transplanted patients are likely to be quite different from those of importance for untransplanted individuals.[33] Individuals transplanted after age l were included in the lth landmark data set and their deaths were counted as events in the survival analysis. The predictors in the lth landmark data set were the three baseline variables, calendar year and variables that summarize the measurements of the remaining 13 time-dependent predictors up to age l. We summarize time-dependent measurements in two ways. Firstly, we used the most recently available measure at time l of each time-dependent variable. This 'last-observation-carried-forward' approach was used in the original descriptions of landmarking.[20,21] Secondly, we fitted a mixed effects model to data available on time-dependent variables up to the landmark age and used the resulting fitted values and slopes at the landmark age as predictors, since some studies have suggested that this makes better use of the data than last-observation-carried-forward.[34–36] We implemented this two-stage landmarking approach by fitting a multivariate mixed model to three continuous time-

ACCEPTED

dependent variables - FEV1%, FVC%, weight - up to each landmark age (eAppendix 2; http://links.lww.com/EDE/B407, eTable 2; http://links.lww.com/EDE/B407).

We created a single stacked data set by stacking the 33 landmark data sets ($l = 18, \ldots, 50$), for use in pooled models (see below). Many individuals appear multiple times in the stacked data set because they are eligible for several landmark data sets. Robust standard errors were used to account for this.

Model building

The aim was to obtain a dynamic prediction model that performs well for predicting 2-, 5- and 10-year survival from each landmark age. We considered a number of multivariable Cox models (Table 2) before selecting a final model based on assessment of their predictive performance. Further details on the models and on how predicted survival probabilities were obtained are given in eAppendix 2; http://links.lww.com/EDE/B407.

Models 1-5 use the last-observation-carried-forward values for the 13 time-dependent predictors. We began by fitting separate survival models from each landmark age l (Model 1). An alternative is to fit a pooled model (a 'supermodel') to the stacked data set. The simplest supermodel (Model 2) allowed a separate baseline hazard for each landmark age, but assumed common predictor coefficients across all landmark ages. Models 1 and 2 were initially fitted using a time horizon of 10 years ($t_{hor} = 10$), which enables us to obtain predicted survival probabilities for any time up to 10 years after the landmark age. We also investigated whether 2- and 5-year survival could be better predicted by using $t_{hor} = 2$ and $t_{hor} = 5$ respectively. One might expect to better predict 2-year survival (for example) by using $t_{hor} = 2$ instead of $t_{hor} = 10$ because the effects of time-dependent variables are expected to change less over 2 years than 10 years. However, this was not found to be the case and all subsequent models were fitted with $t_{hor} = 10$. Since we found that the

ACCEPTED

supermodel gave better predictive performance, subsequently investigated models were all extensions of Model 2.

Model 3 allows predictor coefficients (log hazard ratios) to vary smoothly with l. Model 4 allows the predictor coefficients to vary with time since landmark $(t - l)$. Model 5 uses a common baseline hazard with the impact of landmark age on the hazard modeled using regression terms. Model 6 extends Model 2 by using the fitted value and slope at each landmark age for each of FEV1%, FVC%, and weight from the multivariate mixed models (one for each landmark age) as additional time-dependent predictors (as well as the last-observation-carried-forward values). By incorporating slopes from the mixed models, the prediction model includes information about trajectories of FEV1%, FVC%, and weight up to each landmark age. For height and the categorical time-dependent variables we used last-observation-carried-forward in all models. In all models continuous variables were assumed to have linear effects; modeling them using splines brought negligible changes in predictive performance.

Model assessment

We divided the data into a training-plus-validation set - an 80% random sample of the stacked data, stratified by landmark age - and a "holdout" set - the remaining 20%.[37] The training-plus-validation set was used for model development and assessment. Details are given in eAppendix 3; http://links.lww.com/EDE/B407.

We compared the predictive performances of different models in terms of discrimination, using the C-index,[38–40] and prediction error, using the Brier score.[41,42] C-indexes and Brier scores were calculated separately for each landmark age for prediction of 2-, 5- and 10-year survival. We also obtained overall C-indexes and Brier scores across landmark ages for 2-, 5- and 10-year survival. A Monte-Carlo cross-validation procedure was used to avoid over-optimism about predictive performance.[43]

ACCEPTED

We selected the model with the best predictive performance as the final model, though where several models had similar performance we favored a simpler model. The final model was applied to the holdout data to estimate its performance in a new set of individuals. Last, the final model was fitted to the complete data and is reported in full for use by other researchers. We performed all analyses using R. eAppendix 4; http://links.lww.com/EDE/B407 provides details on software.

RESULTS

Data overview

The stacked data set has 43,592 rows and 6181 unique individuals, of whom 931 died within 10 years of follow-up (eAppendix 2; http://links.lww.com/EDE/B407). Censoring is due to the end of follow-up at the end of 2015, rather than loss to follow-up (eAppendix 2; http://links.lww.com/EDE/B407). Many individuals appear in multiple landmark data sets. eFigure 1; http://links.lww.com/EDE/B407 illustrates how the data arose. Figure 1 summarizes the number of individuals in each landmark data set, and the number of deaths within 2, 5, and 10 years of each landmark age. eTable 1; http://links.lww.com/EDE/B407 gives more detailed information. eTable 3; http://links.lww.com/EDE/B407 summarizes the predictors at landmark ages 20, 30, 40, and 50.

Comparison of dynamic prediction models

Overall C-indexes and Brier scores from Models 1-6 are shown in Table 3. Model 1, in which separate models were fitted from each landmark, gave overall C-Indexes of 0.841 for 2-year survival, 0.811 for 5-year survival, and 0.771 for 10-year survival, and corresponding Brier scores of 0.038 for 2-, 0.082 for 5- and 0.147 for 10-year survival, indicating better predictive performance for short-term survival. A supermodel fitted across landmark ages (Model 2) brought gains in terms of both discrimination (C-indexes) and prediction error (Brier scores). The C-indexes increased to 0.873 for 2-, 0.843 for 5- and 0.804 for 10-year survival, and the

ACCEPTED

Brier scores reduced to 0.036 for 2-, 0.076 for 5-, and 0.133 for 10-year survival. Landmark-age-specific C-indexes and Brier scores (eFigures 2 and 3; http://links.lww.com/EDE/B407) show that the gains in predictive performance from using the supermodel are particularly important for older landmark ages. This is because there are fewer data at those ages and hence more to be gained by drawing strength from other landmark ages by using a supermodel.

Allowing the predictor coefficients to depend on landmark age in a smooth way (Model 3) resulted in very similar results to Model 2. Including time-varying coefficients for all predictors (Model 4) resulted in worse predictive performance compared with Model 2. Restricting the time-varying coefficients to FEV1%, the strongest predictor, gave very similar results to Model 2. Using splines instead of a linear form for the time-varying coefficients did not bring any improvements. This lack of advantage of using time-varying coefficients in part reflects our finding that using a shorter time-horizon ($t_{hor} = 2$ or 5) did not improve prediction. Using a common baseline hazard, with the impact of landmark age modeled using regression terms (Model 5), resulted in considerably worse predictive performance than Model 2.

Inclusion of the fitted values and slopes from mixed models for FEV1%, FVC%, and weight in addition to the last-observation-carried-forward terms brought small improvements in the C-indexes and Brier scores. Further investigations found that including the mixed model terms without the corresponding last-observation-carried-forward terms resulted in worse predictive performance than Models 2 and 6.

Final model

Based on the above comparisons, we selected Model 2 as the final model: increasing model complexity had not resulted in improvements in predictive performance, suggesting a trade-off between increased complexity and estimation of more parameters. While there were small

ACCEPTED

gains in predictive performance from using mixed models for three of the continuous variables (Model 6), these were fairly negligible and came at the expense of a substantially more complicated procedure for obtaining predicted survival probabilities. Also, Model 2 requires only the most recent values of predictors at the landmark age, while the mixed modeling approach (Model 6) requires a series of measures up to the landmark age. Furthermore, Model 2 is more straightforward to explain and report to potential users. eFigure 4; http://links.lww.com/EDE/B407 shows calibration plots for the final model for landmark ages 20, 30, 40, and 50, which compare model-based predicted survival probabilities with 'observed' probabilities. For 2-year and 5-year survival the points lie close to the line of equality, indicating good agreement between predicted probabilities from the model and the observed probabilities. There is also good agreement for 10-year survival for landmark ages 20, 30, and 40. At landmark age 50 the agreement between predicted and observed 10-year survival probabilities is less good, which may be partly due to sparse data at the older ages. These results indicate that the model is well calibrated for prediction of 2- and 5-year survival from all landmark ages, and for 10-year survival at least up to age 40.

Application in the holdout data

The final model was fitted to the complete training-plus-validation data and applied to the holdout data to demonstrate its use in practice. The resulting overall C-indexes were for 0.854 for 2- , 0.843 for 5-, and 0.815 for 10-year survival. The corresponding overall Brier scores were 0.034, 0.077, and 0.125, representing percentage reductions in prediction error against the Kaplan-Meier estimates of survival probabilities of 12.22%, 20.92%, and 23.86%. eTable 4; http://links.lww.com/EDE/B407 summarizes observed survival within groups defined by the predicted survival probabilities.

ACCEPTED

Full model specification

We fitted the final model to the complete data (the training-plus-validation and holdout data combined). Estimated baseline hazards $h_{0l}(t)$ are given in supplementary materials (eAppendix 5; http://links.lww.com/EDE/B407); in combination with the regression coefficients in Table 4, these provide a full specification of the dynamic prediction model. Higher FEV1%, FVC% and weight were strongly associated with reduced hazard. *B. cepacia* infection, CF-related diabetes, and more hospital days on IV antibiotics were strongly associated with increased hazard. Using the final model fitted to the complete data, we calculated 2-, 5- and 10-year predicted survival probabilities from ages 20, 30, 40 and 50 for individuals in the CF Registry at these ages during the most recent 3-year period for which data were available (2013-2015). eFigures 5-8; http://links.lww.com/EDE/B407 and illustrate typical profiles of individuals within groups defined by predicted survival probabilities and show corresponding predicted survivor curves, illustrating in particular how FEV1%, FVC%, weight, CFRD and IV days are associated with survival Figure 2 shows the distributions of the predicted probabilities. At age 20, over 80% of individuals had a greater than 95% probability of 2-year survival, and over 35% of 10-year survival. At landmark ages 30, 40 and 50, over 75% of individuals had a greater than 90% probability to survive 2 years, and over 50% had a greater than 90% probability to survive 5 years. These plots further demonstrate how the model could be used to identify patients at greatest risk and those with a good prognosis.

DISCUSSION

We have developed a model for dynamic prediction of survival for people with CF in the UK using UK CF Registry data. We used a landmarking approach applied to CF data to our knowledge for the first time, making efficient use of the longitudinal data, by using information from the same individual at several ages and incorporating updated measures of

ACCEPTED

health status. The model enables predictions of survival up to 10 years for adults with CF aged up to 50 and can be used to identify high risk patients, making use of information on 16 variables. R code for obtaining estimated survival probabilities from the final model is provided at https://github.com/ruthkeogh/landmark_CF. There are several potential roles for practical use of the model, including for guiding treatment decisions, informing referral for lung transplantation[44], and providing personalized information going far beyond the population-level statistics that are currently available, which is important for patients.

We have outlined a systematic approach to development of a dynamic prediction model using landmarking, incorporating the assessment of models of different levels of complexity by comparing their predictive performance. There have been relatively few practical applications of landmarking.[34,45,46] Unlike previous applications we have provided predicted survival curves instead of focusing on a single time-horizon, and we provided results on model performance for 2-, 5-, and 10-year survival. Prediction of long-term survival is of particular relevance for chronic conditions such as CF, and ours is to our knowledge the first prediction model based on UK CF Registry data. Of the three earlier prediction models using national patient registry data , two used logistic regression,[14,17] and so did not handle censoring, and did not make efficient use of the longitudinal data. Aaron et al.[16] used a stochastic process model. No previous prediction models in CF have considered survival to more than one time point or beyond 5 years.[12–17,22,25] Comparisons of predictive performance with models obtained in other populations are summarized in eAppendix 6; http://links.lww.com/EDE/B407. Future work may result in new models for the UK population that could be compared with ours and it is important that similar measures of predictive performance are presented across studies to facilitate comparisons. We used the landmarking approach to perform dynamic prediction. An alternative approach uses joint modeling of the longitudinal and survival processes.[47–49] Landmarking had several strengths

ACCEPTED

over joint modeling for this application. Firstly, landmarking enabled us to handle transplanted individuals in a straightforward way. We excluded previously transplanted individuals at each landmark age, but retained post-transplant deaths in the data set for estimating survival after each landmark age. Our predictions therefore refer to individuals who are untransplanted at the time of making the prediction. Development of a prediction model for post-transplant survival is an area for further work. It is not clear how transplanted individuals should be handled in the joint modeling approach, especially using readily available software. Secondly, the set of predictors included 12 time-dependent variables of different types (continuous, categorical, binary). Although joint modeling has recently been extended for use with multivariate longitudinal outcomes,[50] its feasibility for use with a large number of such variables of different types remains in question. The two-stage landmarking approach,[34–36] which used mixed models for continuous time-dependent predictors (Model 6), did not result in material gains compared with using the last-observation-carried forward method. Landmarking also has the advantage of being based on methods, notably Cox regression, that are familiar to a clinical audience, which facilitates its explanation. Recent comparisons of landmarking with joint modeling using simulation studies have tended to find joint modeling to perform slightly better than landmarking. [35,36,51] However, they have focused on simple simulation scenarios favouring the joint model and have not considered landmark supermodels.

A major strength of our study is the use of the UK CF Registry data to create the dynamic prediction model. The Registry collects longitudinal data on almost all UK CF patients, and the structured data collection means there is little missing data and little loss to follow-up. A limitation is that predicted survival probabilities cannot account for improvements in survival that are not yet known about, e.g. due to new treatments.[52,53] However, treatments manifest themselves in measures of health status, and so it is likely that the prediction model could still

ACCEPTED

apply. That is, the distribution of health status measures in the CF population may change, but the associations of health status measures with survival remain the same. The standardized format of the Registry data collection means that the model could be assessed and updated if necessary after a few years.

We selected a set of predictors previously associated with survival in CF and collected routinely in the Registry.[3,10] FEV1% is the strongest predictor, though predictive performance is improved by incorporating the additional variables (eTable 5; http://links.lww.com/EDE/B407). Further investigations using variable selection techniques tended to result in a model containing most of the variables. Extensions of variable selection techniques to the context of dynamic prediction remains an area for further methodologic work. There are many other variables in the Registry and an area for further work is to investigate whether using additional variables could improve predictive performance. We took the decision not to use data on treatment use as predictors. As noted above, the impact of treatments on survival is expected to manifest primarily via the health status measures used as predictors. Further investigations also found that adding information on use of two treatments did not materially improve prediction (eTable 5; http://links.lww.com/EDE/B407). Furthermore, the models created in this work are designed with prediction in mind and the estimated coefficients associated with the predictor variables do not necessarily represent causal effects. Inclusion of treatment variables could create danger of misinterpretation of the impacts of treatment on survival prediction curves as causal effects, which could result in inappropriate withholding of treatment if treatment is (non-causally) associated with worse prognosis. Estimation of treatment effects using patient registry data is an area of growing interest, [54,55] but involves a separate question from that focused on in this paper.

Our model is for adults with CF. There are relatively few deaths in CF patients aged under 18 in the UK and different variables may be important for survival prediction in children.[12,56]

ACCEPTED

We restricted to predictions for adults aged up to 50 because the data above age 50 are sparse. Investigations into the health of older people with CF are of interest.

In summary, we have developed a novel landmarking model for dynamic prediction of survival for people with CF in the UK. Further work involves the practical implementation of our model in a form suitable for use by clinicians, potentially as an add-on to patient information that can already be viewed via the Registry interface. In addition, it is important that patients and caregivers are supported to interpret personalized survival predictions.[57–59]

ACCEPTED

References

1. Cystic Fibrosis Trust. *UK Cystic Fibrosis Registry Annual data report 2016*. (2017).
2. Rowe, S. M., Miller, S. & Sorscher, E. J. Cystic Fibrosis. *N. Engl. J. Med.* **352,** 1992–2001 (2005).
3. MacNeill, S. J. Epidemiology of Cystic Fibrosis. in *Hodson and Geddes' Cystic Fibrosis* (eds. Bush, A., Bilton, D. & Hodson, M.) 18–40 (CRC Press, 2016).
4. Lim, M. *et al.* Diagnosis of cystic fibrosis in London and South East England before and after the introduction of newborn screening. *Arch Dis Child* **99,** 197–202 (2014).
5. Dodge, J. A., Lewis, P. A., Stanton, M. & Wilsher, J. Cystic fibrosis mortality and survival in the UK: 1947-2003. *Eur. Respir. J.* **29,** 522–526 (2007).
6. Elborn, J. S., Shale, D. J. & Britton, J. R. Cystic fibrosis: current survival and population estimates to the year 2000. *Thorax* **46,** 881–885 (1991).
7. Keogh, R. H. & Stanojevic, S. A guide to interpreting estimated median age of survival in cystic fibrosis patient registry reports. *J. Cyst. Fibros.* **17,** 213–217 (2018).
8. Goss, C., Quon, B. & Bilton, D. Quality improvement: using databases to improve cystic fibrosis care. in *Hodson and Geddes' Cystic Fibrosis* (eds. Bush, A., Bilton, D. & Hodson, M.) 425–431 (CRC Press, 2016).
9. Jackson, A. D. & Goss, C. H. Epidemiology of CF: How registries can be used to advance our understanding of the CF population. *J. Cyst. Fibros.* (2017). doi:10.1016/j.jcf.2017.11.013
10. Buzzetti, R. *et al.* An overview of international literature from cystic fibrosis registries: 1. Mortality and survival studies in cystic fibrosis. *J. Cyst. Fibros.* **8,** 229–237 (2009).
11. George, P. M. *et al.* Improved survival at low lung function in cystic fibrosis: cohort study from 1990 to 2007. *BMJ* **342,** d1008 (2011).
12. Aurora, P., Wade, A., Whitmore, P. & Whitehead, B. A model for predicting life

ACCEPTED

expectancy of children with cystic fibrosis. *Eur. Respir. J.* **16,** 1056–1060 (2000).

13. Hayllar, K. M. *et al.* A prognostic model for the prediction of survival in cystic fibrosis. *Thorax* **52,** 313–7 (1997).

14. Liou, T. G. *et al.* Predictive 5-year survivorship model of cystic fibrosis. *Am. J. Epidemiol.* **153,** 345–352 (2001).

15. Mayer-Hamblett, N., Rosenfeld, M., Emerson, J., Goss, C. H. & Aitken, M. L. Developing Cystic Fibrosis Lung Transplant Referral Criteria Using Predictors of 2-Year Mortality. *Am. J. Respir. Crit. Care Med.* **166,** 1550–1555 (2002).

16. Aaron, S. D., Stephenson, A. L., Cameron, D. W. & Whitmore, G. A. A statistical model to predict one-year risk of death in patients with cystic fibrosis. *J. Clin. Epidemiol.* **68,** 1336–1345 (2015).

17. Nkam, L. *et al.* A 3-year prognostic score for adults with cystic fibrosis. *J. Cyst. Fibros.* **119,** 682–689 (2017).

18. Keogh, R. H., Szczesniak, R., Taylor-Robinson, D. & Bilton, D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data. *J. Cyst. Fibros.* **17,** 218–227 (2018).

19. Taylor-Robinson, D. *et al.* Data Resource Profile: The UK Cystic Fibrosis Registry. *Int. J. Epidemiol.* **47,** 9–10e (2017).

20. Van Houwelingen, H. C. Dynamic Prediction by Landmarking in Event History Analysis. **34,** 70–85 (2007).

21. van Houwelingen, H. C. & Putter, H. *Dynamic prediction in clinical survival analysis*. (CRC Press/Chapman and Hall, 2012).

22. McCarthy, C., Dimitrov, B. D., Meurling, I. J., Gunaratnam, C. & McElvaney, N. G. The CF-ABLE score: A novel clinical prediction rule for prognosis in patients with

ACCEPTED

cystic fibrosis. *Chest* **143,** 1358–1364 (2013).

23. Courtney, J. M. *et al.* Predictors of mortality in adults with cystic fibrosis. *Pediatr. Pulmonol.* **42,** 525–532 (2007).

24. Milla, C. E. & Warwick, W. J. Risk of Death in Cystic Fibrosis Patients With Severely Compromised Lung Function. *Chest* **113,** 1230–1234 (1998).

25. Kerem, E., Reisman, J., Corey, M., Canny, G. J. & Levison, H. Prediction of Mortality in Patients with Cystic Fibrosis. *N. Engl. J. Med.* **326,** 1187–1191 (1992).

26. Keating, C. *et al.* Reduced survival in adult cystic fibrosis despite attenuated lung function decline. *J. Cyst. Fibros.* **16,** 78–84 (2016).

27. Stephenson, A. L. *et al.* A contemporary survival analysis of individuals with cystic fibrosis: A cohort study. *Eur. Respir. J.* **45,** 670–679 (2015).

28. Hulzebos, E. H. J. *et al.* Prediction of mortality in adolescents with cystic fibrosis. *Med. Sci. Sports Exerc.* **46,** 2047–2052 (2014).

29. Quanjer, P. H. *et al.* Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur. Respir. J.* **40,** 1324–43 (2012).

30. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19,** 716–723 (1974).

31. Flume, P. A. & Vandevanter, D. R. Pulmonary exacerbations. in *Hodson and Geddes' Cystic Fibrosis* (eds. Bush, A., Bilton, D. & Hodson, M.) 221–235 (CRC Press, 2016).

32. Bhatt, J. M. Treatment of pulmonary exacerbations in cystic fibrosis. *Eur. Respir. Rev.* **22,** 205–16 (2013).

33. Stephenson, A. L. *et al.* Clinical and demographic factors associated with post-lung transplantation survival in individuals with cystic fibrosis. *J. Hear. Lung Transplant.* **34,** 1139–1145 (2015).

34. Paige, E. *et al.* Landmark Models for Optimizing the Use of Repeated Measurements

ACCEPTED

of Risk Factors in Electronic Health Records to Predict Future Disease Risk. *Am. J. Epidemiol.* **kwy018,** (2018).

35. Rizopoulos, D., Molenberghs, G. & Lesaffre, E. M. E. H. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical J.* **59,** 1261–1276 (2017).

36. Maziarz, M., Heagerty, P., Cai, T. & Zheng, Y. On longitudinal prediction with time-to-event outcome: Comparison of modeling options. *Biometrics* **73,** 83–93 (2016).

37. Yong, F., Cai, T., Wei, L. & Tian, L. Classical model selection. in *Handbook of Survival Analysis* (eds. Klein, J., van Houwelingen, H., Ibrahim, J. & Scheike, T.) (CRC Press, 2014).

38. Uno, H., Cai, T., Pencina, M. J., D'Agostino, R. B. & Wei, L. J. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat. Med.* **30,** 1105–1117 (2011).

39. Pencina, M. J., D'Agostino, R. B., D'Agostino, R. B. & Vasan, R. S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat. Med.* **27,** 157–172 (2008).

40. Gerds, T. A., Kattan, M. W., Schumacher, M. & Yu, C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Stat. Med.* **32,** 2173–2184 (2013).

41. Graf, E., Schmoor, C., Sauerbrei, W. & Schumacher, M. Assessment and comparison of prognostic classification schemes for survival data. *Stat. Med.* **18,** 2529–2545 (1999).

42. Gerds, T. A. & Schumacher, M. Consistent Estimation of the Expected Brier Score in General Survival Models with Right-Censored Event Times. *Biometrical J.* **48,** 1029–1040 (2006).

ACCEPTED

43. Kuhn, M. & Johnson, K. Over-Fitting and Model Tuning. in *Applied Predictive Modeling* 61–92 (Springer, 2013).

44. NHS Blood and Transplant Cardiothoracic Advisory Group. *POLICY POL231 / 2 Lung Candidate Selection Criteria*. (2017).

45. Lehners, N. *et al.* Analysis of long-term survival in multiple myeloma after first-line autologous stem cell transplantation: impact of clinical risk factors and sustained response. *Cancer Med.* **7,** 307–316 (2018).

46. Fontein, D. B. Y. *et al.* Dynamic prediction in breast cancer: proving feasibility in clinical practice using the TEAM trial. *Ann. Oncol.* **26,** 1254–1262 (2015).

47. Tsiatis, A. A. & Davidian, M. Joint modeling of Longitudinal and Time-to-Event Data: An Overview. *Stat Sin.* **14,** 809–834 (2004).

48. Rizopoulos, D. *Joint models for longitudinal and time-to event data with applications in R*. (CRC Press/Chapman and Hall, 2012).

49. Asar, Z. R., Ritchie, J., Kalra, P. A. & Diggle, P. J. Joint modeling of repeated measurement and time-to-event data: an introductory tutorial. *Int. J. Epidemiol.* **44,** 334–344 (2015).

50. Rizopoulos, D. & Ghosh, P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat. Med.* **30,** 1366–1380 (2011).

51. Suresh, K., Taylor, J. M. G., Spratt, D. E., Daignault, S. & Tsodikov, A. Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model. *Biometrical J.* **59,** 1277–1300 (2017).

52. Solomon, G. M., Marshall, S. G., Ramsey, B. W. & Rowe, S. M. Breakthrough therapies: Cystic fibrosis (CF) potentiators and correctors. *Pediatr. Pulmonol.* **50 Suppl 4,** S3–S13 (2015).

53. Murphy, M. P. & Caraher, E. Current and Emerging Therapies for the Treatment of

ACCEPTED

Cystic Fibrosis or Mitigation of Its Symptoms. *Drugs R. D.* **16,** 1–17 (2016).

54. Newsome, S., Keogh, R. & Daniel, R. Estimating long-term treatment effects in observational data: a comparison of the performance of different methods under real-world uncertainty. *Stat. Med.* (2018).

55. Sawicki, G. S. *et al.* Sustained Benefit from Ivacaftor Demonstrated by Combining Clinical Trial and Cystic Fibrosis Patient Registry Data. *Am. J. Respir. Crit. Care Med.* **192,** 836–842 (2015).

56. McColley, S. A. *et al.* Risk factors for mortality before age 18 years in cystic fibrosis. *Pediatr. Pulmonol.* **52,** 909–915 (2017).

57. O'Connor, A. M., Légaré, F. & Stacey, D. Risk communication in practice: the contribution of decision aids. *BMJ* **327,** 736–40 (2003).

58. Rakow, T., Wright, R. J., Bull, C. & Spiegelhalter, D. J. Simple and multistate survival curves: Can people learn to use them? *Med. Decis. Mak.* **32,** 792–804 (2012).

59. Spiegelhalter, D. Risk and Uncertainty Communication. *Annu. Rev. Stat. Its Appl.* **4,** 31–60 (2017).

ACCEPTED

Figure Legends

Figure 1. Overview of number of individuals in each landmark data set. On the left: Number of individuals alive at each landmark age at any point during the study period. On the right: Number of deaths within 2-, 5-, and 10-years after each landmark age, among those alive at each landmark age.

Figure 2. Plots showing the distribution of 2-, 5- and 10-year survival probabilities from landmark ages 20, 30, 40, and 50 for individuals in the Registry at those ages between 2013 and 2015. [This plot is shown in color in eFigure 9; http://links.lww.com/EDE/B407].

ACCEPTED

Table 1. Variables considered as predictors. All are time-dependent except the baseline variables.

Variable category	Variables	Description	Further information
Baseline variables	Sex	Male (0), Female (1)	
	Genotype	F508del: Homozygous F508del: Heterozygous F508del: No copies	
	Age of diagnosis	In years.	
Calendar year	Calendar year	2005-2015 (coded as 0-10)	
Lung function	FEV1%	FEV1% predicted, obtained using GLI equations.	Measured at the annual review visit.
	FVC%	FVC% predicted, obtained using GLI equations.	
Height and weight	Weight	Kilograms (kg)	Measured at the annual review visit.
	Height	Centimetres (cm)	
Microbiology	*Pseudomonas aeruginosa*	No (0), Yes (1)	Any finding based on microbiology results since the last annual review.
	Burkholderia cepacia	No (0), Yes (1)	
	Staphylococcus aureus	No (0), Yes (1)	
	Methicillin-resistant *Staphylococcus aureus (MRSA)*	No (0), Yes (1)	
Complications	Pancreatic insufficiency[a]	No (0), Yes (1)	All -in the year prior to the annual review.
	CF related diabetes[a]	No (0), Yes (1)	
	Number of hospital IV days[b]	0 days (reference category) 1-14 days 15-28 days 29+ days	
	Number of home IV days[b]	0 days (reference category) 1-14 days 15-28 days 29+ days	
	Hospitalisation (not for IVs)	No (0), Yes (1)	

FEV1%: Percent predicted forced expiratory volume in 1 second.
FVC%: Percent predicted forced vital capacity.
IV: Intravenous antibiotic therapy.
[a] Once an individual was recorded as being pancreatic insufficient ("Yes" (1)) they were considered to be pancreatic insufficient at all subsequent time points. Once an individual was recorded as having CFRD ("Yes" (1)) they were considered to have CFRD at all subsequent time points.
[b] Number of hospital and home IV days are used as surrogate indicators of pulmonary exacerbations.

ACCEPTED

Table 2. Summary of dynamic prediction models investigated. In all analyses the timescale is age (t). Landmark age is denoted l. For models 1 and 2, using age as the time scale or time-since-landmark as the timescale are exactly equivalent.

Model	Form of the log hazard: $\log h_l(t\|X(l), X^*(l), Z)$	Description
Model 1	$\log h_{0l}(t) + \beta_l^T X(l) + \gamma_l^T Z, l = 1, \dots, L$	Separate model fitted at each landmark age
Model 2	$\log h_{0l}(t) + \beta^T X(l) + \gamma^T Z$	Supermodel with separate baseline hazards for $l = 1, \dots, L$ and common predictor coefficients across landmark ages.
Model 3	$\log h_{0l}(t) + \beta(l)^T X(l) + \gamma(l)^T Z$	Supermodel with separate baseline hazards for $l = 1, \dots, L$ and predictor coefficients modelled as a function of landmark age l.
Model 4	$\log h_{0l}(t) + \beta(t-l)^T X(l) + \gamma(t-l)^T Z$	Supermodel with separate baseline hazards for $l = 1, \dots, L$ and time-varying predictor coefficients, but common across landmark ages.
Model 5	$\log h_0(t) + \beta^T X(l) + \gamma^T Z + f(l;\delta)$	Supermodel with an overall baseline hazard, common predictor coefficients across landmark ages, and landmark effects $f(l;\delta)$.
Model 6	$\log h_{0l}(t) + \beta^T X(l) + \gamma^T Z + \theta^T X^*(l)$	As in Model 2, but with additional predictors $X^*(l)$ from the multivariate mixed model.

$h_l(t|X(l), X^*(l), Z)$: Hazard at time t given $X(l)$, Z and $X^*(l)$, and given eligibility for the lth landmark data set (Supplementary Section S1).
$h_{0l}(t)$: Baseline hazard at time t given eligibility for the lth landmark data set (Supplementary Section S1)..
Z: Vector of baseline predictors (sex, genotype and age of diagnosis).
$X(l)$: Vector of the last-observation carried forward values at landmark age l for time-dependent predictors (calendar year, FEV1%, FVC%, weight, height, CFRD. pancreatic insufficiency, *P. aeruginosa*, *B. cepacia*, *S. aureus*, MRSA, non-IV hospitalization, number of IV days).
$X^*(l)$: Vector of predicted values and slopes for FEV1%, FVC% and weight from a multivariate mixed model.
FEV1%: Percent predicted forced expiratory volume in 1 second.
FVC%: Percent predicted forced vital capacity.
IV: Intravenous antibiotic therapy.
MRSA: Methicillin-resistant *Staphylococcus aureus*.

ACCEPTED

Table 3. Overall C-Indexes, Brier scores, and Brier score percentage reductions[a] for prediction of 2-year, 5-year and 10-year survival from Models 1-6.

	C-Index			Brier score			Brier score % reduction[a]		
	2-year	5-year	10-year	2-year	5-year	10-year	2-year	5-year	10-year
Model 1	0.841	0.811	0.771	0.038	0.082	0.147	9.56	15.54	11.67
Model 2	0.873	0.843	0.804	0.036	0.076	0.133	14.85	21.79	20.58
Model 3	0.872	0.843	0.803	0.036	0.076	0.132	14.798	22.32	21.14
Model 4[b]	0.837	0.837	0.797	0.043	0.088	0.168	-2.29	9.85	-0.70
Model 4[c]	0.873	0.843	0.804	0.036	0.076	0.133	14.68	21.61	20.09
Model 5	0.849	0.813	0.766	0.039	0.087	0.158	7.53	11.00	5.57
Model 6	0.873	0.844	0.805	0.036	0.076	0.132	14.73	21.84	20.91

Model 1: separate landmark models
Model 2: supermodel with common β coefficients across landmarks and separate baseline hazard for each landmark age
Model 3: supermodel with interactions between each covariate and $, l$ and separate baseline hazard for each landmark age
Model 4: supermodel with time-varying β coefficients and separate baseline hazard for each landmark age
Model 5: supermodel with common β coefficients across landmarks, overall baseline hazard, and landmark effects
Model 6: as in Model 2, with the addition of mixed model terms to the predictors.
a Percentage reduction in the Brier score relative to the Brier score obtained from Kaplan-Meier estimates of survival probabilities (fitted separately from each landmark age with no predictors).
b Including time-varying coefficients for all variables.
c Including time-varying coefficients for FEV1% only.

ACCEPTED

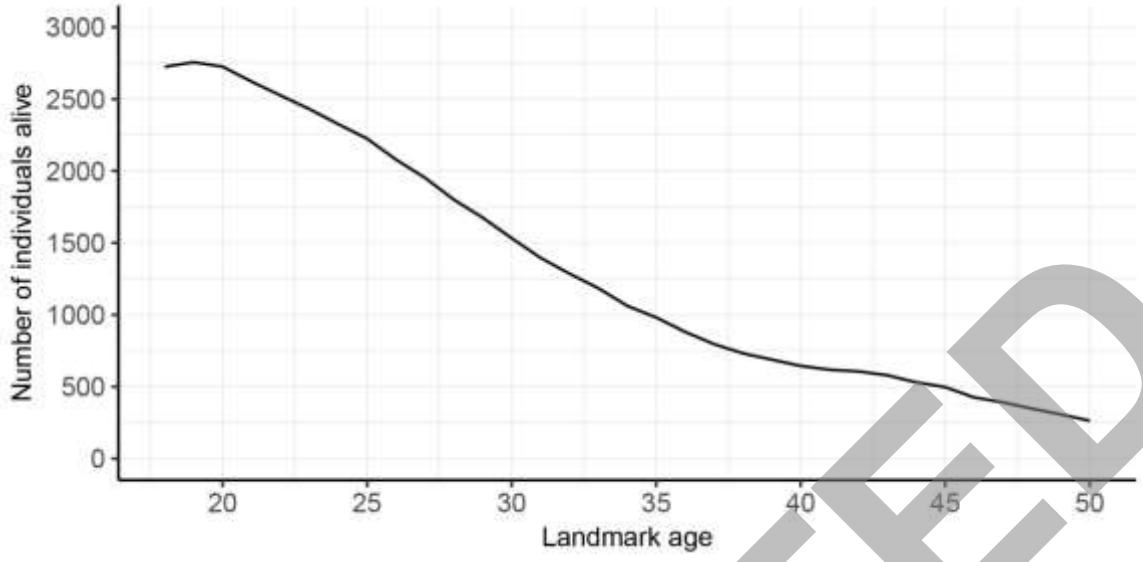
Table 4. Results from fitting the final selected model to the complete data. HR: hazard ratio. CI: confidence interval. The confidence intervals were obtained using robust standard errors. HRs for continuous variables refer to a unit change.

Variable		HR	95% CI
Sex	Male	1 (ref)	
	Female	0.87	(0.72,1.06)
Genotype	2 copies	1 (ref)	
	1 copy	0.98	(0.83,1.15)
	Other	1.05	(0.78,1.43)
Age of diagnosis (years)		0.99	(0.98,1.00)
Calendar year (years)		0.97	(0.95,1.00)
FEV1%		0.97	(0.96,0.97)
FVC%		0.99	(0.98,1.00)
Weight (kg)		0.98	(0.97,0.99)
Height (cm)		0.99	(0.98,1.00)
P. aeruginosa	No	1 (ref)	
	Yes	1.04	(0.90,1.19)
B. cepacia	No	1 (ref)	
	Yes	1.91	(1.51,2.40)
S. aureus	No	1 (ref)	
	Yes	0.87	(0.77,0.98)
MRSA	No	1 (ref)	
	Yes	1.02	(0.77,1.34)
Pancreatic insufficiency	No	1 (ref)	
	Yes	1.07	(0.80,1.42)
CF related diabetes	No	1 (ref)	
	Yes	1.48	(1.29,1.70)
Hospitalisation (not for IVs)	No	1 (ref)	
	Yes	1.06	(0.79,1.41)
Number of hospital IV days	0 days	1 (ref)	
	1-14 days	1.13	(0.99,1.28)
	15-28 days	1.52	(1.31,1.76)
	29+ days	2.37	(2.05,2.74)
Number of home IV days	0 days	1 (ref)	
	1-14 days	1.03	(0.90,1.19)
	15-28 days	1.06	(0.90,1.26)
	29+ days	1.39	(1.20,1.61)

FEV1%: Percent predicted forced expiratory volume in 1 second.
FVC%: Percent predicted forced vital capacity.
IV: Intravenous antibiotic therapy.
MRSA: Methicillin-resistant *Staphylococcus aureus*

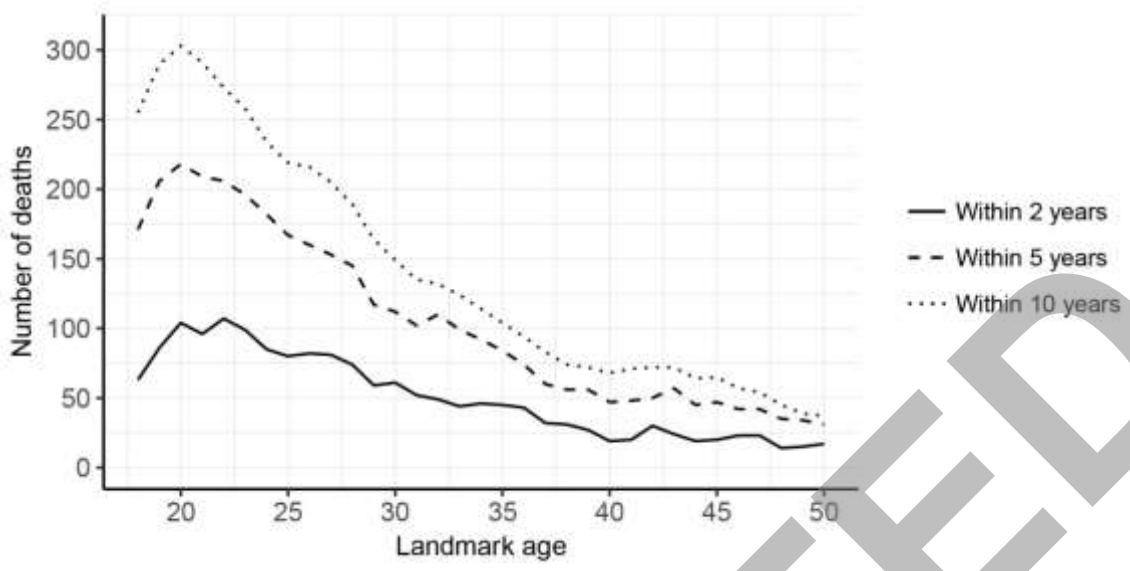
ACCEPTED

Figure 1a



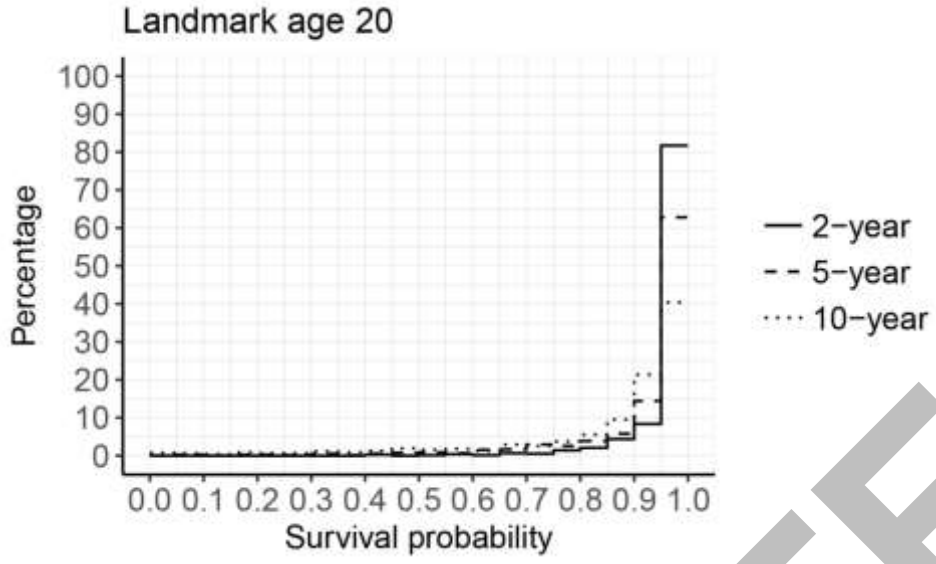
ACCEPTED

Figure 1b



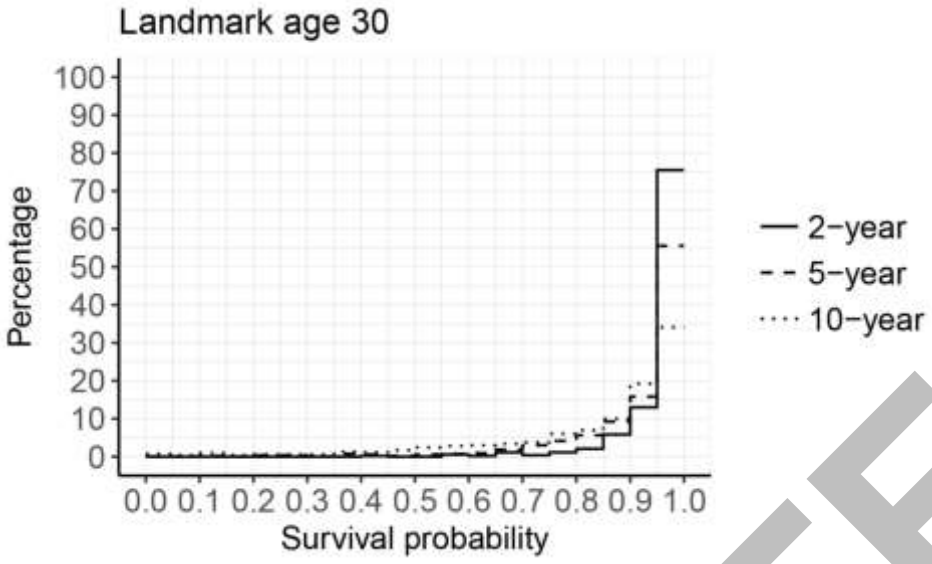
ACCEPTED

Figure 2a



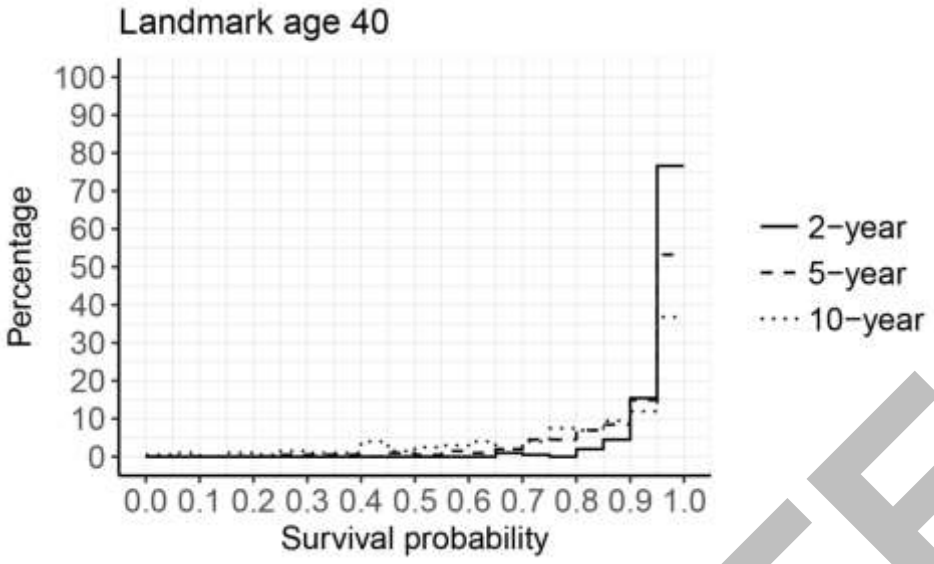
ACCEPTED

Figure 2b



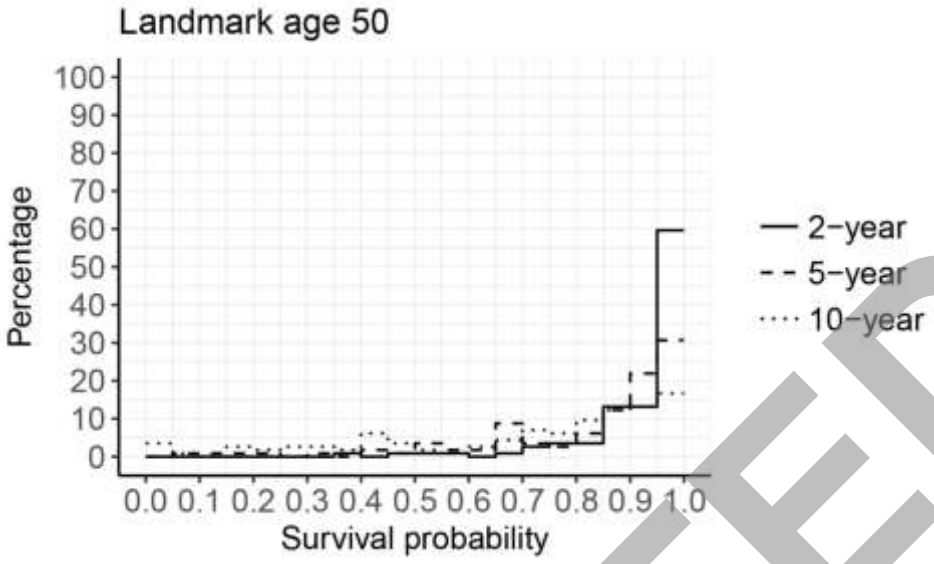
ACCEPTED

Figure 2c



ACCEPTED

Figure 2d



ACCEPTED