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.

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.

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

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 personalised estimates of survival probabilities given a patient's current health status using 16 predictors. The model was developed using the landmarking approach, giving predicted survival curves up to 10 years from ages 18 to 50. Several models were compared 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 personalised prediction models can guide treatment decisions and provide personalised 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.

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

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 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 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 personalised prediction of survival in the UK making use of time-dependent measures of health status.

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.[3,10,11,13,15–17,22–28] This set consists of 3 baseline variables – sex; genotype (F508del alleles); age of diagnosis - 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 (CFRD); pancreatic insufficiency; days in hospital on intravenous (IV) antibiotics; days at home on IV antibiotics; and other hospitalisation. FEV1% and FVC% were calculated 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. Individuals who died or were lost-to-follow-up before 2005 were therefore excluded. 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. two years after the 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 $l = 18, \ldots, 50$ (Supplementary Figure 1, Supplementary Section S1). 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 summarise the measurements of the remaining 13 time-dependent predictors up to age l. We summarise 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' (LOCF) 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 LOCF.[34–36] We implemented this two-stage landmarking approach by fitting a multivariate mixed model to three continuous time-dependent variables - FEV1%, FVC%, weight - up to each landmark age (Supplementary Section S2, Supplementary Table 2).

A single stacked data set was created, 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 Supplementary Section S2.

Models 1-5 use the LOCF 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 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 modelled 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 LOCF 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 LOCF in all models. In all models continuous variables were assumed to have linear effects; modelling them using splines brought negligible changes in predictive performance.

Model assessment

The data were divided into a "training+validation" (TV) set - an 80% random sample of the stacked data, stratified by landmark age - and a "holdout" set - the remaining 20%.[37] The TV set was used for model development and assessment. Details are given in Supplementary Section S3.

The predictive performances of different models were compared 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]

The final model was selected as that with the best predictive performance, though where several models had similar performance we favoured a simpler model. The final model was applied to the holdout data to estimate its performance in a new set of individuals. Lastly, the final model was fitted to the complete data and is reported in full for use by other researchers.

All analyses were performed using R. Supplementary Section S4 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 (Supplementary Section S2). Censoring is entirely due to the end of follow-up at the end of 2015, rather than loss-to-follow-up. Many individuals appear in multiple landmark data sets. Supplementary Figure 1 illustrates how the data arose. Figure 1 summarises the number of individuals in each landmark data set, and the number of deaths within 2, 5 and 10 years of each landmark age. Supplementary Table 1 gives more detailed information. Table 3 summarises 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 4. Model 1, in which separate models were fitted from each landmark, gave overall C-Indexes of 0.841, 0.811 and 0.771 for 2-, 5- and 10-year survival respectively, and corresponding Brier scores of 0.038, 0.082 and 0.147, indicating better predictive performance for short-term survival. A supermodel fitted across landmark ages (Model 2) brought gains in terms of both discrimination and prediction error. The C-indexes for 2-, 5- and 10-year survival increased to 0.873, 0.843 and 0.804, and the Brier scores reduced to 0.036, 0.076, and 0.133. Landmark-age-specific C-indexes and Brier scores (Supplementary Figures 2 and 3) show that the gains in predictive performance from using the supermodel are particularly important for older landmark ages. This is because there are less 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 modelled 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 LOCF terms brought small improvements in the C-indexes and Brier scores. Further investigations found that including the mixed model terms without the corresponding LOCF 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 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 significantly 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 modelling 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.

Figure 2 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 TV data and applied to the holdout data to demonstrate its use in practice. The resulting overall C-indexes for 2-, 5- and 10-year survival were 0.854, 0.843, and 0.815. 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%. Supplementary Table 3 summarises observed survival within groups defined by the predicted survival probabilities.

Full model specification

The final model was fitted to the complete data (the TV and holdout data combined). Estimated baseline hazards $h_{0l}(t)$ are given in Supplementary Materials (Section S5); in combination with the regression coefficients in Table 5, these provide a full specification of the dynamic prediction model. Higher FEV1%, FVC% and weight were strongly associated with reduced hazard. *B. cepacia* infection, CFRD, and more hospital IV days 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). Figure 3 and Supplementary Figures 4-6 illustrate typical profiles of individuals within groups defined by predicted survival probabilities and show corresponding predicted survivor curves. Figure 4 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 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 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. There are several potential roles for practical use of the model, including for guiding treatment decisions, informing referral for lung transplantation[44], and providing personalised information going far beyond the population-level statistics that are currently available, which is important for patients (Keogh 2017, unpublished manuscript).

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 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 summarised in Supplementary Section S 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 modelling of the longitudinal and survival processes.[47–49] Landmarking had several strengths over joint modelling 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 modelling approach, especially using readily available software. Secondly, the set of predictors included 12 time-dependent variables of different types (continuous, categorical, binary). Although joint modelling 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 LOCF 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 modelling using simulation studies have tended to find joint modelling 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 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 (Supplementary Table 4). 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 methodological 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 (Supplementary Table 4). 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] 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 care-givers are supported to interpret personalised survival predictions.[57–59]

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)	

[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.

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 LOCF 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.

Table 3. Descriptive statistics at landmark ages 20, 30, 40 and 50. Summaries are given as the number (N) and percent for categorical variables and as median and interquartile range (IQR) for continuous variables.

Variable		Landmark age 20		Landmark age 30		Landmark age 40		Landmark age 50	
		N	%	N	%	N	%	N	%
Sex	Male	1443	52.9	863	56.3	385	59.8	160	60.8
	Female	1283	47.1	670	43.7	259	40.2	103	39.2
Genotype	2 copies	1549	56.8	820	53.5	263	40.8	87	33.1
	1 copy	956	35.1	567	37.0	312	48.4	141	53.6
	Other	221	8.1	146	9.5	69	10.7	35	13.3
Age of diagnosis (years)	Median (IQR)	0.3	(0.1, 2.0)	0.7	(0.1, 3.5)	2.0	(0.3, 18.1)	13.0	(1.0, 36.0)
Calendar year	Median (IQR)	2010	(2008, 2013)	2011	(2009, 2013)	2011	(2008, 2013)	2012	(2009, 2014)
FEV1%	Median (IQR)	69.4	(52.1, 85.6)	60.5	(42.8, 78.6)	55.3	(38.1, 74.7)	53.9	(36.6, 72.3)
FVC%	Median (IQR)	83.0	(68.6, 95.8)	79.9	(63.4, 92.3)	77.6	(61.2, 91.3)	74.7	(62.6, 89.6)
Weight (kg)	Median (IQR)	57.0	(50.4, 65.3)	63.0	(55.3, 72.2)	66.1	(58.9, 75.6)	69.0	(60.5, 79.5)
Height (cm)	Median (IQR)	166.3	(160.0, 173.1)	169.0	(162.0, 176.0)	169.0	(162.9, 175.0)	169.5	(162.0, 176.0)
P. aeruginosa	No	1127	41.3	471	30.7	234	36.3	107	40.7
	Yes	1599	58.7	1062	69.3	410	63.7	156	59.3
B. cepacia	No	2621	96.1	1445	94.3	604	93.8	253	96.2
	Yes	105	3.9	88	5.7	40	6.2	10	3.8
S. aureus	No	1580	58.0	940	61.3	410	63.7	167	63.5
	Yes	1146	42.0	593	38.7	234	36.3	96	36.5
MRSA	No	2651	97.2	1480	96.5	628	97.5	255	97.0
	Yes	75	2.8	53	3.5	16	2.5	8	3.0
Pancreatic insufficiency	No	224	8.2	189	12.3	150	23.3	87	33.1
	Yes	2502	91.8	1344	87.7	494	76.7	176	66.9
CF related diabetes	No	1968	72.2	914	59.6	382	59.3	158	60.1
	Yes	758	27.8	619	40.4	262	40.7	105	39.9
Hospitalisation (not for IVs)	No	2649	97.2	1483	96.7	626	97.2	250	95.1
	Yes	77	2.8	50	3.3	18	2.8	13	4.9
Number of hospital IV days	0 days	1648	60.5	958	62.5	458	71.1	187	71.1
	1-14 days	487	17.9	274	17.9	109	16.9	37	14.1
	15-28 days	245	9.0	125	8.2	36	5.6	19	7.2
	29+ days	346	12.7	176	11.5	41	6.4	20	7.6
Number of home IV days	0 days	1852	67.9	931	60.7	425	66.0	188	71.5
	1-14 days	340	12.5	227	14.8	85	13.2	28	10.6
	15-28 days	229	8.4	132	8.6	50	7.8	20	7.6
	29+ days	305	11.2	243	15.9	84	13.0	27	10.3

Table 4. 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.

Table 5. Results from fitting the final selected model to the complete data. HR: hazard ratio. CI: confidence interval. The confidence intervals and p-values were obtained using robust standard errors.

Variable		HR	95% CI	P-value
Sex	Male	1 (ref)		
	Female	0.87	(0.72,1.06)	0.16
Genotype	2 copies	1 (ref)		
	1 copy	0.98	(0.83,1.15)	0.78
	Other	1.05	(0.78,1.43)	0.74
Age of diagnosis		0.99	(0.98,1.00)	0.17
Calendar year		0.97	(0.95,1.00)	0.03
FEV1%		0.97	(0.96,0.97)	<0.001
FVC%		0.99	(0.98,1.00)	<0.001
Weight (kg)		0.98	(0.97,0.99)	<0.001
Height (cm)		0.99	(0.98,1.00)	0.17
P. aeruginosa	No	1 (ref)		
	Yes	1.04	(0.90,1.19)	0.63
B. cepacia	No	1 (ref)		
	Yes	1.91	(1.51,2.40)	<0.001
S. aureus	No	1 (ref)		
	Yes	0.87	(0.77,0.98)	0.02
MRSA	No	1 (ref)		
	Yes	1.02	(0.77,1.34)	0.90
Pancreatic insufficiency	No	1 (ref)		
	Yes	1.07	(0.80,1.42)	0.65
CF related diabetes	No	1 (ref)		
	Yes	1.48	(1.29,1.70)	<0.001
Hospitalisation (not for IVs)	No	1 (ref)		
	Yes	1.06	(0.79,1.41)	0.71
Number of hospital IV days	0 days	1 (ref)		
	1-14 days	1.13	(0.99,1.28)	0.07
	15-28 days	1.52	(1.31,1.76)	<0.001
	29+ days	2.37	(2.05,2.74)	<0.001
Number of home IV days	0 days	1 (ref)		
	1-14 days	1.03	(0.90,1.19)	0.66
	15-28 days	1.06	(0.90,1.26)	0.47
	29+ days	1.39	(1.20,1.61)	<0.001

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.

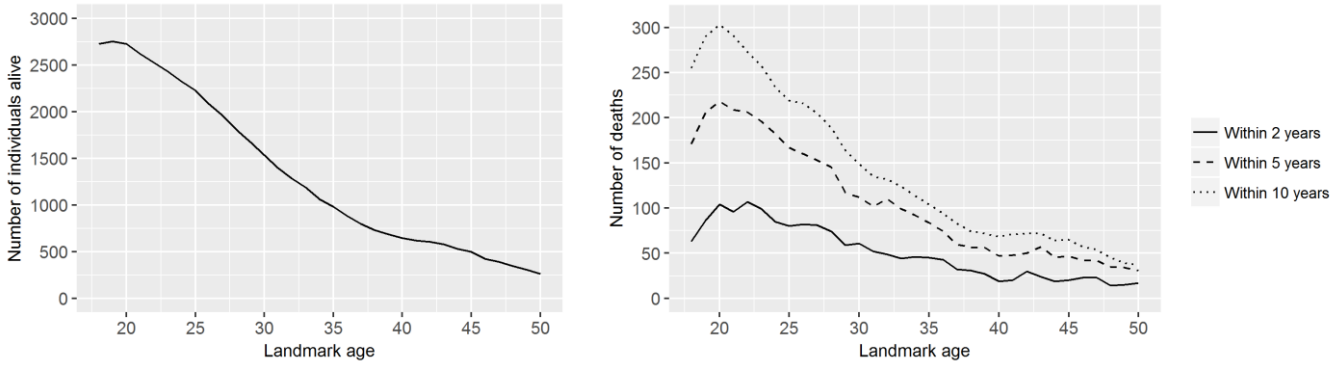


Figure 2. Calibration plots using the final model (Model 2) for prediction of 2-year, 5-year and 10-year survival from landmark ages 20, 30, 40 and 50. The vertical axis shows the mean model-based x-year survival probability (x=2,5,10) in quintiles of the model-based probabilities. The horizontal axis shows the mean x-year survival probability obtained using Kaplan-Meier estimates in quintiles of the model-based probabilities. The five points have been joined by a line. [This plot is shown in colour in Supplementary Figure 5].

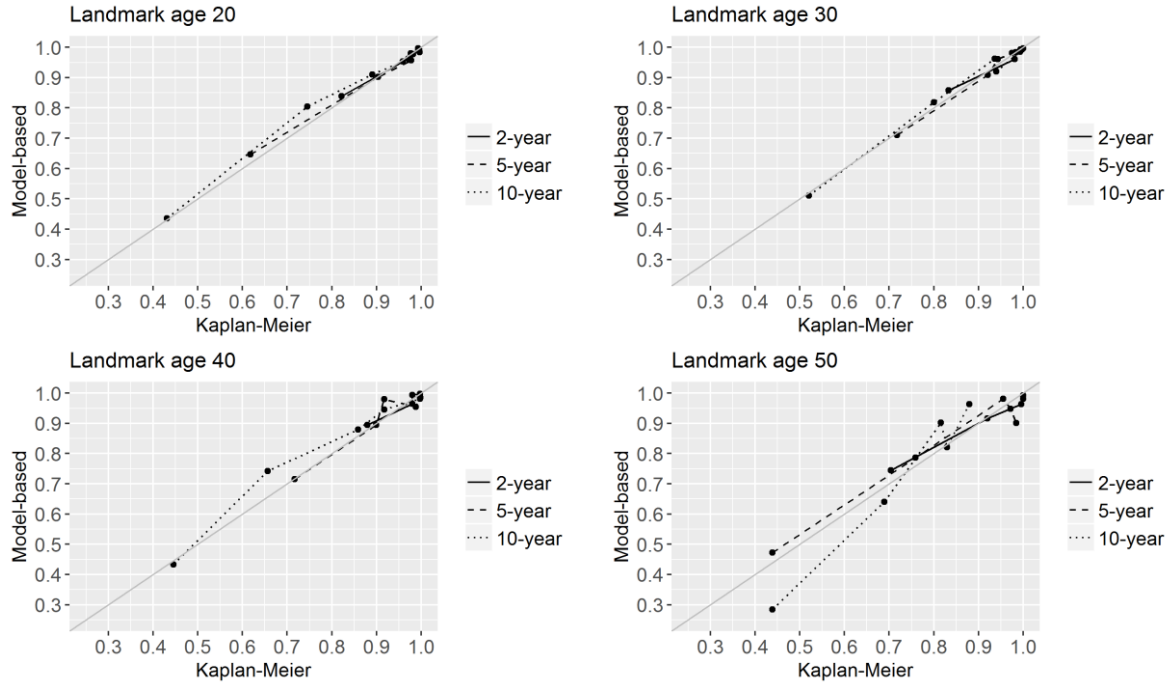


Figure 3. Predicted survival curves from landmark age 30 for example individuals in groups defined by 5-year survival probabilities. For individuals in the Registry at age 30 between 2013 and 2015 we obtained their predicted 5-year survival probabilities and categorized into groups with 5-year survival probabilities <0.5, (0.5,0.7,], (0.7,0.8], (0.8,0.9], (0.9,0.95], (0.95,0.99], (0.99, 1]. An example individual was created for each group. Corresponding results for landmark ages 20, 40 and 50 are shown in Supplementary Figure 4

(i) Characteristics of example individuals[a] in groups defined by 5-year survival probability.

5-year survival probability group	<0.5	(0.5,0.7]	(0.7,0.8]	(0.8,0.9]	(0.9,0.95]	(0.95,0.99]	(0.99,1]
Example person	1	2	3	4	5	6	7
Males/Females[b]							
Genotype (no. copies of F508del)	2	2	2	2	2	2	1
Age of diagnosis (years)	0, 0	0, 0	0, 1	0, 1	0, 0	0, 1	5, 3
FEV1%	29, 25	24, 38	35, 32	36, 43	51, 54	71, 76	91, 97
FVC%	31, 36	49, 57	57, 50	63, 61	69, 70	88, 89	100, 102
Weight (kg)	48, 48	64, 47	60, 52	62, 55	69, 56	70, 58	77, 68
Height (cm)	170, 156	172, 156	173, 156	172, 162	173, 163	174, 162	178, 166
P. aeruginosa	No, Yes	Yes	Yes	Yes	Yes	Yes	Yes
B. cepacia	Yes, No	No	No	No	No	No	No
S. aureus	No	No	No	No	No	No	No, Yes
MRSA	No	No	No	No	No	No	No
Pancreatic insufficiency	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CF related diabetes	Yes	Yes	Yes	Yes	Yes	No	No
Hospitalisation (not for IVs)	No	No	No	No	No	No	No
Number of hospital IV days	29+	29+	29+, 0	0	0	0	0
Number of hospital IV days	0, 29+	0, 29+	0, 29+	0	0	0	0

[a]We created an example individual for each group using the median values of the continuous predictors and the most common value of each categorical variable within that group. This was done separately for males and females.

[b] Values are shown as 'male, female', except were the value for males and females was the same.

(ii) Predicted survivor curves based on the final model for example individuals with characteristics shown in the table above.

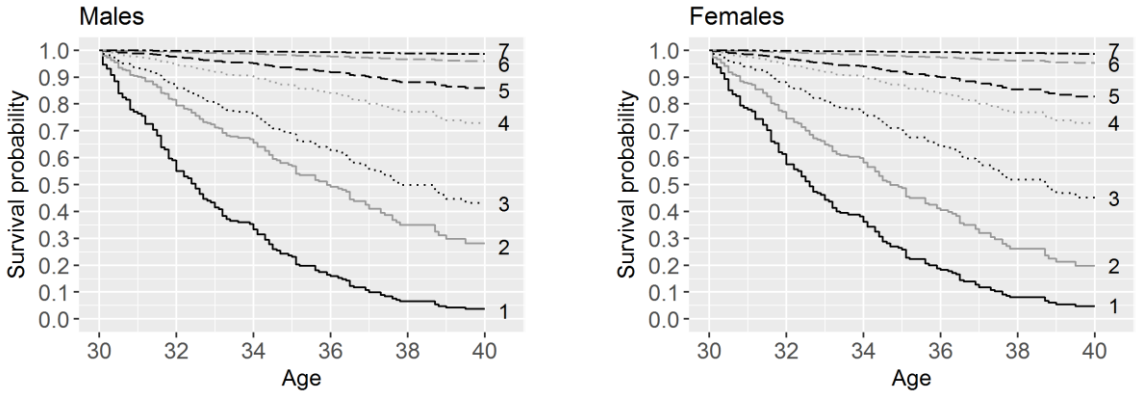
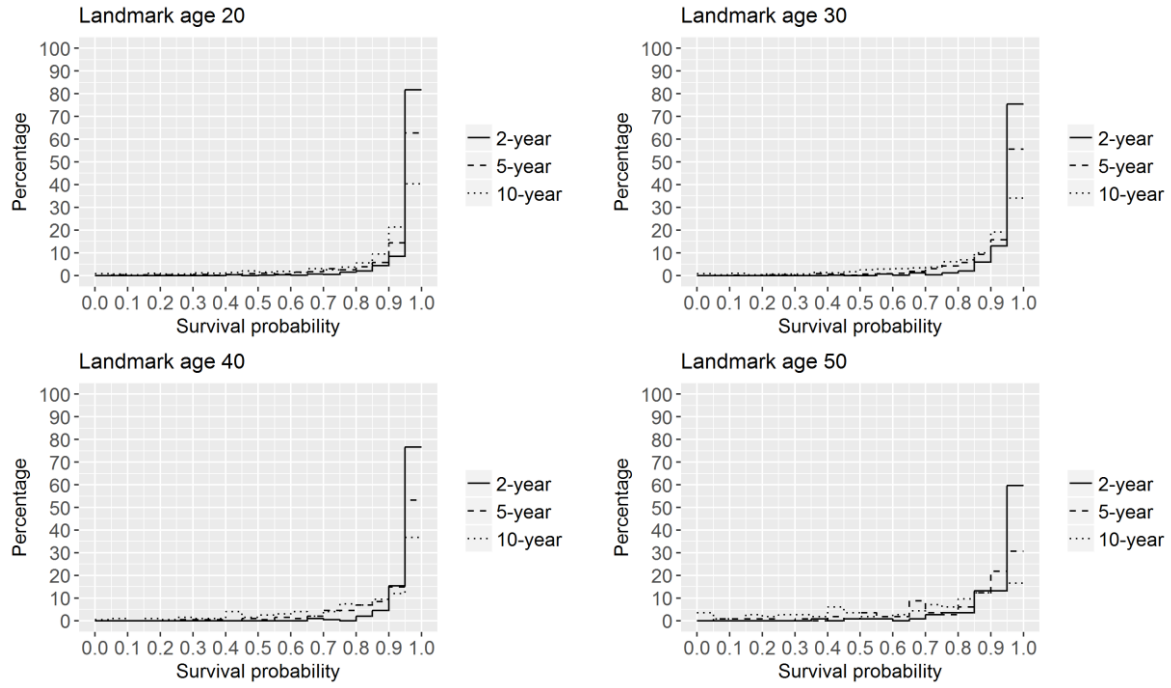


Figure 4. Plots showing the distribution of 2-, 5- and 10-year survival probabilities from landmark ages 20, 30, 40 and 50 for individuals in the Registry at those ages between 2013 and 2015. [This plot is shown in colour in Supplementary Figure 6].



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 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 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 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).

43. Kuhn, M. & Johnson, K. Over-Fitting and Model Tuning. in *Applied Predictive Modelling* 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 modelling 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 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).

Supplementary materials for:

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

S1. Creation of landmark data sets

Supplementary Figure 1 illustrates how the landmark data sets arose. An individual was included in the landmark data set at age l if they met *all* of the following criteria:

- They reached age l between 1st January 2005 and 31st December 2015.
- They joined the Registry prior to reaching age l. The date of joining the Registry is the date of the first annual review at which data were obtained.
- They were diagnosed with CF prior to reaching age l.
- They have not received an organ transplant of any type prior to reaching age l.
- They have measures of all time-dependent variables recorded prior to reaching age l.

We refer to an individual as "eligible for the lth landmark data set" if she/he satisfied these five conditions. Supplementary Table 1 summarises the landmark data sets in terms of number of individuals, number of deaths within 2, 5 and 10 years of the landmark age, and number of censorings.

S2. Survival prediction models

Time scale and follow-up

In all models the time origin is date of birth and analyses are performed using left-truncation at the landmark age. The censoring time was the earliest of death, 31st December 2015 and a specified time horizon t_{hor}. Since dates of birth and death were only available in month/year format, the day was imputed as the 15th of the month. For example, an individual aged 18 on 1st January 2005 (who has been diagnosed, joined the Registry, and not received a transplant) contributes up to 11 years of follow-up until the end of 2015 to the landmark data set for age 18 and up to 10 years of follow-up for the landmark dataset for age 19 (if they do not die, become lost-to-follow-up, or have a transplant between ages 18 and 19), and so on. An individual aged 18 on 1st January 2014 contributes up to 2 years of follow-up to the landmark data set for age 18 and up to 1 year of follow-up for the landmark dataset for age 19.

The UK CF Registry aims to capture deaths from all causes. Of the 931 deaths used in this study, 775 (83.2%) were due to respiratory or cardiorespiratory failure, 55 (5.9%) were transplantation-related, 13 (1.4%) were due to liver disease or failure, 9 (1.0%) were due to cancer, 9 (1.0%) were due to trauma or suicide, 34 (3.7%) were due to "other causes" (recorded in a separate field and including "End state cystic fibrosis" and "Haemoptysis"), 35 (3.9%) were due to an unknown cause, and for 1 individual the cause was not recorded.

We assumed that all deaths are captured and the main results presented assume censoring is entirely administrative. In a sensitivity analysis we treated individuals not recorded at an annual follow-up for over 2 years as lost-to-follow-up. This did not materially alter the results – the C-indexes for 2-5- and 10-year survival from the final model (Model 2) were 0.874, 0.847, 0.807 respectively, and corresponding Brier scores were 0.036, 0.075, 0.130.

Landmark survival models

We let Z denote the vector of baseline predictors (sex, genotype and age of diagnosis) and $X(l)$ denote the vector of the last-observation-carried-forward (LOCF) values for time-dependent predictors (calendar year, FEV%, FEV%, weight, height, CFRD. pancreatic insufficiency, *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Staphylococcus aureus*, *Methicillin-resistant Staphylococcus aureus* (MRSA), non-IV hospitalization, number of IV days) at landmark age l.

Model 1 for the log conditional hazard is

$$\log h_l(t|X(l), Z) = \log h_{0l}(t) + \beta_l^T X(l) + \gamma_l^T Z, l = 1, \dots, L \qquad \text{Model 1}$$

where $h_{0l}(t)$ is the baseline hazard at age t conditional on eligibility for the lth landmark data set, and β_l and γ_l are vectors of log hazard ratios specific to landmark age l. Model 1 is in fact L models, which are fitted in each landmark data set $l = 1, \dots, L$.

Model 2 for the log conditional hazard is

$$\log h_l(t|X(l), Z) = \log h_{0l}(t) + \beta^T X(l) + \gamma^T Z \qquad \text{Model 2}$$

where $h_{0l}(t)$ is again the baseline hazard at age t conditional on eligibility for the lth landmark data set ($l = 1, \dots, L$). β and γ are vectors of log hazard ratios, which are assumed to be the same for all l. Model 2 therefore allows a separate baseline hazard from each landmark age, but common predictor coefficients across all landmark ages. It is fitted in the stacked data set using Cox regression with a stratified baseline hazard.[1,2] We note that for Models 1 and 2, using age as the time scale or time-since-landmark as the timescale are exactly equivalent.

Models 1 and 2 make the proportional hazards assumption that the association of the predictors $X(l)$ and Z with the hazard is the same over time since l, i.e. that the β_l and β parameters are not time-dependent. 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. We also investigated whether 2-year and 5-year survival could be better predicted by using a shorter time horizon by fitting Models 1 and 2 using $t_{hor} = 2$ and $t_{hor} = 5$ respectively.

Model 3 extends Model 2 by allowing the log hazard ratios to depend on l in a smooth way:

$$\log h_l(t|X(l), Z) = \log h_{0l}(t) + \beta(l)^T X(l) + \gamma(l)^T Z \qquad \text{Model 3}$$

where $\beta(l)$ and $\gamma(l)$ denote vectors of log hazard ratios that are functions of l. We considered linear forms $\beta(l) = \beta_0 + \beta \times (l - 18)$ and $\gamma(l) = \gamma_0 + \gamma \times (l - 18)$ and restricted cubic spline forms with knots at 18, 30, 40 and 50. The results reported in Table 4 of the main text are from the analysis using the linear form for $\beta(l)$, as using restricted cubic splines did not materially improve predictive performance.

In Model 4 the supermodel was extended to allow time-varying coefficients, with the association between the predictors and the hazard dependent on time-since landmark ($t - l$):

$$\log h_l(t|X(l), Z) = \log h_{0l}(t) + \beta(t - l)^T X(l) + \gamma(t - l)^T Z \qquad \text{Model 4}$$

where $\beta(t - l)$ and $\gamma(t - l)$ denote vectors of log hazard ratios that are functions of $t - l$. We considered linear forms $\beta(t - l) = \beta_0 + \beta \times (t - l)$ and $\gamma(t - l) = \gamma_0 + \gamma \times (t - l)$ and restricted cubic spline forms with knots at $t - l = 2,5,8$. The results reported in Table 4 of the main text are from the analysis using the linear form for $\beta(t - l)$, as using restricted cubic splines did not materially improve predictive performance.

Model 5 uses an overall baseline hazard instead of separate baseline hazards for each landmark age, with the impact of landmark age modelled using regression terms:

$$\log h_l(t|X(l), Z) = \log h_0(t) + \beta^T X(l) + \gamma^T Z + f(l; \delta) \qquad \text{Model 5}$$

where $h_0(t)$ is a common baseline hazard and $f(l;\delta)$ is a function of landmark age. We used a restricted cubic spline form for $f(l;\delta)$ with knots at 18, 30, 40 and 50.

In Model 6 we extended Model 2 by adding the fitted values and slopes from the multivariate mixed model (see below) for FEV%, FVC% and weight to the set of time-dependent predictors at each landmark age:

$$\log h_l(t|X(l),Z) = \log h_{0l}(t) + \beta^T X(l) + \gamma^T Z + \theta^T X^*(l) \qquad \text{Model 6}$$

where $X^*(l)$ denotes the vector of predicted values and slopes for FEV%, FVC% and weight from the multivariate mixed model.

All models were fitted by maximum partial likelihood.

Multivariate mixed model

A multivariate linear mixed model for FEV1%, FVC%, BMI and weight was fitted to the repeated measures up to landmark age l $(l = 1, \dots, L)$ for individuals in the landmark data set at age l. Separate models were fitted for each landmark age. The longitudinal variables were modelled as a linear function of age with a random intercept and slope. We also included fixed effects of all the other predictors, including both baseline and time-dependent predictors. For each individual in landmark dataset l $(l = 1, \dots, L)$ the individual fitted values and slopes for FEV1%, FVC% and weight at age l were obtained. The numbers of longitudinal measurements used in the multivariate mixed models are summarised in Supplementary Table 2.

Predicted survival probabilities

From each model the predicted survival probability to time t after the landmark age, conditional on survival to the landmark age, on baseline variables Z and on values of time-dependent predictors at the landmark age $X(l)$, $S(l+t|X(l),Z,T>l)$, was obtained using the relationship

$$S(l+t|X(l),Z,T>l) = \exp\left\{-\int_l^{l+t} h(u|X(l),Z,u>l)\,du\right\}$$

For models without time-varying hazard ratios (Models 1-3 and 5-6) we used the estimator:

$$\hat{S}(l+t|X(l),Z,T>l) = \exp\left\{-e^{\hat{\beta}^T X(l)+\hat{\gamma}^T Z}\sum_{l<u\le l+t}\hat{h}_{0u}\right\}$$

where $\hat{h}_{0u}$ denotes the baseline hazard at time u estimated from the increments in Breslow's estimate of the cumulative baseline hazard and the sum is over event times.[3] For Model 4, which has time-varying hazard ratios, we used the estimator

$$\hat{S}(l+t|X(l),Z,T>l) = \exp\left\{-\sum_{l<u\le l+t}\hat{h}_{0u}e^{\hat{\beta}^T(u-l)X(l)+\hat{\gamma}^T(u-1)Z}\right\}$$

S3. Model assessment

Overview

Models were assessed and compared based on the "3-in-1" procedure described by Yong et al (2013), which incorporates model building using cross-validation, final model choice, and statistical inference.[4] The data were first divided into a "training+validation" (TV) set and a "holdout" set. The TV set is used in the model development and assessment. The holdout set is reserved for applying the selected model at the end. No models are fitted using the holdout data. The TV set is a sample of 80% from the stacked

data, stratified by landmark age. The holdout set is formed from the remaining 20% of individuals at each landmark age. Some individuals appear in both the TV and holdout stacked data sets, but not with the same landmark age.

For model assessment we used the C-index,[5–8] the Brier score,[9,10] and percentage reduction in the Brier score relative to the null model (i.e. the model excluding all predictors, using Kaplan-Meier estimates).[11] The C-index and Brier scores were obtained using inverse probability of censoring weights. For Model 4 we accommodated the time-varying coefficients into the estimation of the C-Index and Brier score.[8]

A Monte-Carlo cross-validation procedure was used within the TV data set to avoid over-optimism due to overfitting [12]. The procedure was as follows:

(i) An 80% stratified random sample, with stratification by landmark age l, was obtained from the TV data set.
(ii) The model was fitted on the 80% sample.
(iii) The fitted model was used to obtain predicted survival probabilities to a given time from each landmark age l (see below) for the 20% not in the sample.
(iv) Model performance measures (C-index, Brier score, and percentage reduction in the Brier score) were obtained in the 20% not in the sample on which the model was fitted.
(v) Steps (i)-(iv) were repeated 200 times and we obtained the average C-index, Brier score and Brier score reduction across the 200 samples.

Model assessment measures were obtained for 2-year, 5-year and 10-year survival from each landmark age. Therefore there are 99 averaged C-indices and Brier scores for each model (33×3, where 33 is the number of landmark ages 18-50). For each model we also obtained an overall C-index and Brier score which are not age-adjusted. Further details are given below. To simplify the notation we give the details of the C-index and Brier score as if applied to the complete stacked data (the TV and holdout data combined).

Truncated C-Index

The following description of the C-index follows that of Gerds et al..[7] Let T_i and C_i denote respectively the event time and censoring time for individual i. We observe $\tilde{T}_i = \min(T_i, C_i)$ and the event indicator $\Delta_i = 1(T_i < C_i)$. Let $\hat{S}_l(l+t|X(l), Z)$ denote the estimated probability of survival beyond age $l + t$ conditional on survival to age l and given predictor values $X(l), Z$ at age l. The truncated C-index is

$$C_l(t) = E_{ij}\{1\{\hat{S}_l(l+t|X_i(l), Z_i) < \hat{S}_l\big(l+t|X_j(l), Z_i\big)\}|T_i < T_j, T_i \le l+t, T_i > l, T_j > l\}$$

where the expectation is with respect to two subjects i, j, both alive at age l ($T_i > l$). Not all pairs of individuals i, j are comparable. We can compare two individuals who both have the event prior to age $l + t$; two individuals, one of whom has the event prior to age $l + t$ and the other of which is known to be alive (censored) at age $l + t$. We cannot compare two individuals who are both known to be alive (censored) at age $l + t$, two individuals both censored before age $l + t$, or a pair in which one individual has the event and the other is censored before the other's event time. The fact that not all pairs of individuals can be compared is handled using inverse probability of censoring weights (IPCW). The truncated C-Index can be expressed as

$$C_l(t)$$
$$= \frac{E_{ij}\{1\{\hat{S}_l(l+t|X_i(l), Z) < \hat{S}_l\big(l+t|X_j(l), Z_j\big)\}|T_i > l, T_j > l\} E_{ij}\{T_i < T_j, T_i \le l+t|T_i > l, T_j > l, X_i(l), X_j(l)\}}{\Pr\big(T_i < T_j, T_i \le l+t|T_i > l, T_j > l\big)}$$
$$= \frac{E_{ij}\left\{1\{\hat{S}_l(l+t|X_i(l), Z_i) < \hat{S}_l\big(l+t|X_j(l), Z_i\big)\} \int_l^{l+t} S\big(u|X_j(l), Z_i, u > l\big)\, S(du|X_i(l), Z, u > l)|T_i > l, T_j > l\right\}}{E_{ij}\left\{\int_l^{l+t} S\big(u|X_j(l), Z, u > l\big)\, S(u|X_i(l), Z, u > l)\right\}}$$

We assume that the event and censoring time are independent conditional on the variables, i.e. $C_i \coprod T_i | X_i(l), T_i > l, C_i > l$, and that the probability of being uncensored at the prediction horizon $l + t$ is bounded away from 0. This gives rise to the IPCW estimator

$$\hat{C}_l(t) = \frac{\sum_{i=1}^{n_l}\sum_{j=1}^{n_l} 1\{\hat{S}_l(l+t|X_i(l),Z) < \hat{S}_l(l+t|X_j(l),Z)\}1\{\tilde{T}_i < \tilde{T}_j\}1\{\tilde{T}_i \le l+t, \Delta_i = 1\}\widehat{W}_{ij}^{-1}}{\sum_{i=1}^{n_l}\sum_{j=1}^{n_l} 1\{\tilde{T}_i < \tilde{T}_j\}1\{\tilde{T}_i \le l+t, \Delta_i = 1\}\widehat{W}_{ij}^{-1}}$$

of $C_l(t)$, where $\widehat{W}_{ij} = \widehat{\Pr}(C_j > \tilde{T}_i | X_j(l), Z, \tilde{T}_j > l)\,\widehat{\Pr}(C_i \ge \tilde{T}_i | X_i(l), Z, \tilde{T}_i > l)$ is a weight, where the censoring probabilities used in the weight are obtained from a model to be specified (see below).

The C-index $C_l(t)$ is conditional on survival to age l and a separate estimated C-index is obtained for any combination of l and t ($l = 18, \ldots, 50; t = 2,5,10$). We also considered an overall C-index which is combined across landmark ages. Consider the stacked landmark data set and let L_i denote the landmark age for record (row) i. Some individuals appear in more than one row in the stacked landmark data set and we define $ID(i)$ to be the unique identifier (ID number) for the individual in row i. The overall C-index is

$$C_{\text{overall}}(t) = E_{ij}\Big\{1(ID(i) \ne ID(j))1\Big\{\hat{S}_{L_i}(L_i + t|X_i(l),Z) < \hat{S}_{L_j}\big(L_j + t|X_j(l),Z\big)\Big\}|(T_i - L_i) < (T_j - L_i), (T_i - L_i) \le t\Big\}$$

where the expectation is with respect to two rows i, j in the stacked landmark data set. Inclusion of the indicator $1(ID(i) \ne ID(j))$ ensures that an individual is not compared with herself/himself. An estimator incorporating censoring weights is

$$\hat{C}_{\text{overall}}(t) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} 1\Big\{\hat{S}_{L_i}(L_i + t|X_i(l),Z) < \hat{S}_{L_j}\big(L_j + t|X_j(l),Z\big)\Big\}1\{\widetilde{(T_i - L_i)} < (\tilde{T}_j - L_i)\}1\{\tilde{T}_i - L_i \le t, \Delta_i = 1\}\widehat{W}_{ij}^{*-1}}{\sum_{i=1}^{N}\sum_{j=1}^{N} 1\{(\tilde{T}_i - L_i) < (\tilde{T}_j - L_j\}1\{(\tilde{T}_i - L_i) \le t, \Delta_i = 1\}\widehat{W}_{ij}^{*-1}}$$

where N is the total number of individuals in the stacked landmark data set and the weights are $\widehat{W}_{ij}^{*} = \widehat{\Pr}((C_j - L_j) > (\tilde{T}_i - L_i)|X_j(L_j), \tilde{T}_j > L_j)\,\widehat{\Pr}((C_i - L_i) \ge (\tilde{T}_i - L_i)|X_i(L_i), \tilde{T}_i > L_i)$.

We assumed that the probabilities in the weights $\widehat{W}_{ij}$ do not depend on $X_j(l)$ or Z and therefore used $\widehat{\Pr}(C_j > \tilde{T}_i | \tilde{T}_j > l)$ in place of s $\widehat{\Pr}(C_j > \tilde{T}_i | X_j(l), \tilde{T}_j > l)$ and $\widehat{\Pr}(C_i \ge \tilde{T}_i | \tilde{T}_i > l)$ in place of $\widehat{\Pr}(C_i \ge \tilde{T}_i | X_i(l), Z, \tilde{T}_i > l)$. The probabilities were estimated separately from each landmark age using Kaplan-Meier estimates. A similar approach was used for the weights $\widehat{W}_{ij}^{*}$.

In summary we obtained $\hat{C}_{\text{overall}}(t)$ for $t = 2,5,10$ and $\hat{C}_l(t)$ for $t = 2,5,10$ and $l = 18, \ldots, 50$.

Brier score

The Brier score is the mean squared prediction error. As for the C-index, we obtained separate Brier scores at each landmark age and an overall brier score. In the absence of censoring an estimator of the Brier score is

$$\hat{B}_l(t) = \frac{1}{n_l}\sum\nolimits_{i \in D_l}\big\{\hat{S}_{il}(l + t|X_i(l), Z_i) - I_i(T_i > l + t|T_i > l)\big\}^2$$

where $\hat{S}_{il}(l + t|X_i(l), Z_i)$ is the model-based estimated probability of survival to age $l + t$ for individual i in the landmark data set at age l, $I_i(t > l + t|T_i > l)$ is the observed indicator of survival to age $l + t$, and the sum is over the n_l individuals in landmark data set l (D_l). An estimator incorporating inverse probability of censoring weights is

$$\hat{B}_l(t) = \frac{1}{n_l}\sum_{i\in D_l} I(d_i = 1 \cup T_i > l + t)\left\{\hat{S}_{il}(l + t|X_i(l), Z_i) - I_i\left(T_i > l + t\middle|\tilde{T}_i > l\right)\right\}^2 \widehat{W}_i^{-1}$$

where d_i is the event indicator, $I(d_i = 1 \cup T_i > l + t)$ is an indicator taking value 1 for individuals who have the event or whose censoring age is after $t + l$, and zero otherwise, and $\widehat{W}_i = \widehat{\Pr}(C_i > \min(T_i^-, l + t)|\tilde{T}_i > l)$ is the probability of being censored beyond age $\min(T_i^-, l + t)$. The inverse probability of censoring weights were obtained using Kaplan-Meier estimates stratified by landmark age.

The overall Brier score estimator is

$$\hat{B}_{\text{overall}}(t) = \frac{1}{N}\sum_i I(d_i = 1 \cup T_i > L_i + t)\left\{\hat{S}_{iL_i}(L_i + t|X_i(L_i), Z_i) - I_i\left(T_i > L_i + t\middle|\tilde{T}_i > l\right)\right\}^2 \widehat{W}_i^{-1}$$

where the sum is over all rows in the stacked landmark data set and $\widehat{W}_i = \widehat{\Pr}(C_i > \min(T_i^-, L_i + t)|\tilde{T}_i > L_i)$.

Brier scores were also obtained under a null model using Kaplan-Meier estimates of the survival probabilities stratified by landmark age but with no other predictors. These are denoted $\hat{B}_{l,\text{null}}(t)$ and $\hat{B}_{\text{overall,null}}(t)$. The percentage reduction in the Brier score from a given model compared with the null model was calculated using $100(\hat{B}_{l,\text{null}}(t) - \hat{B}_l(t))/\hat{B}_{l,\text{null}}(t)$ and $100(\hat{B}_{\text{overall,null}}(t) - \hat{B}_{\text{overall}}(t))/\hat{B}_{\text{overall,null}}(t)$.

In summary we obtained $\hat{B}_{\text{overall}}(t)$ for $t = 2{,}5{,}10$ and $\hat{B}_l(t)$ for $t = 2{,}5{,}10$ and $l = 18, \ldots, 50$, and the corresponding percentages reductions in the Brier score relative to the null model.

Calibration plots

After selecting the final model, calibration plots were obtained to show graphically the agreement between predicted survival probabilities from the model and the 'true' probabilities. The steps for creating these plots were as follows:

Steps (i)-(iii) are the same as described earlier, in the Overview section of S3.
(iv) The predicted 2-year survival probabilities from landmark age l were divided into quintiles and we obtained the mean predicted 2-year survival probability for individuals within each quintile, denoted $\bar{S}(2)_{l,Q1}, \bar{S}(2)_{l,Q2}, \bar{S}(2)_{l,Q3}, \bar{S}(2)_{l,Q4}, \bar{S}(2)_{l,Q5}$. We also obtained the Kaplan-Meier estimate of 2-year survival for the individuals within each quintile, denoted $KM(2)_{l,Q1}, KM(2)_{l,Q2}, KM(2)_{l,Q3}, KM(2)_{l,Q4}, KM(2)_{l,Q5}$. The same was done for 5-year and 10-year survival.
(v) Steps (i)-(iv) were repeated 200 times and for each $l = 18, \ldots, 50$ we obtained the average of each $\bar{S}(2)_{l,Q1}, \ldots, \bar{S}(2)_{l,Q5}$ and the average of each $KM(2)_{l,Q1}, K \ldots, KM(2)_{l,Q5}$ across the 200 samples..
(vi) The averaged $\bar{S}(2)_{l,Q1}, \ldots, \bar{S}(2)_{l,Q5}$ from step (v) were plotted against the averaged $KM(2)_{l,Q1}, K \ldots, KM(2)_{l,Q5}$.

Calibration plots for landmark ages 20,30,40 and 50 are shown in the main text Figure 2. In a well-calibrated model the five points lie on the $y = x$ line.

S4. Software

All analyses were performed using R. The landmark models described in Section S2 can be fitted easily using the coxph function from the survival package after some rearrangement of the data.[13] Some of the data rearrangement can be performed using the dynpred package,[14] for example using the cutLM function, though we did not use that here. Estimated survival probabilities can be obtained using

'predict' after coxph, though special code was written to obtain the predicted survival probabilities from Model 4, which included time-varying coefficients.

There exist various packages for obtaining C-indexes and Brier scores. None of the existing functions for estimating the C-index appear to accommodate a stratified baseline hazard, and so we used bespoke code. We used 'pew' from the dynpred package to estimate the Brier scores; this requires pre-estimation of matrices of predicted survival and censoring probabilities.

The multivariate mixed model used to obtain the additional predictors $X^*(l)$ for Model 6 was fitted using the lme function from the nlme package.[15] Existing software, including the nlme package, does not appear to allow out-of-sample predictions from mixed models. We therefore used bespoke code which is available from https://github.com/ruthkeogh/landmark_CF.

S5. Final model specification

Example code for obtaining estimated survival probabilities from the final model is provided at https://github.com/ruthkeogh/landmark_CF. This includes csv files containing estimated cumulative baseline hazards for each landmark age ($l = 18, \ldots, 50$).

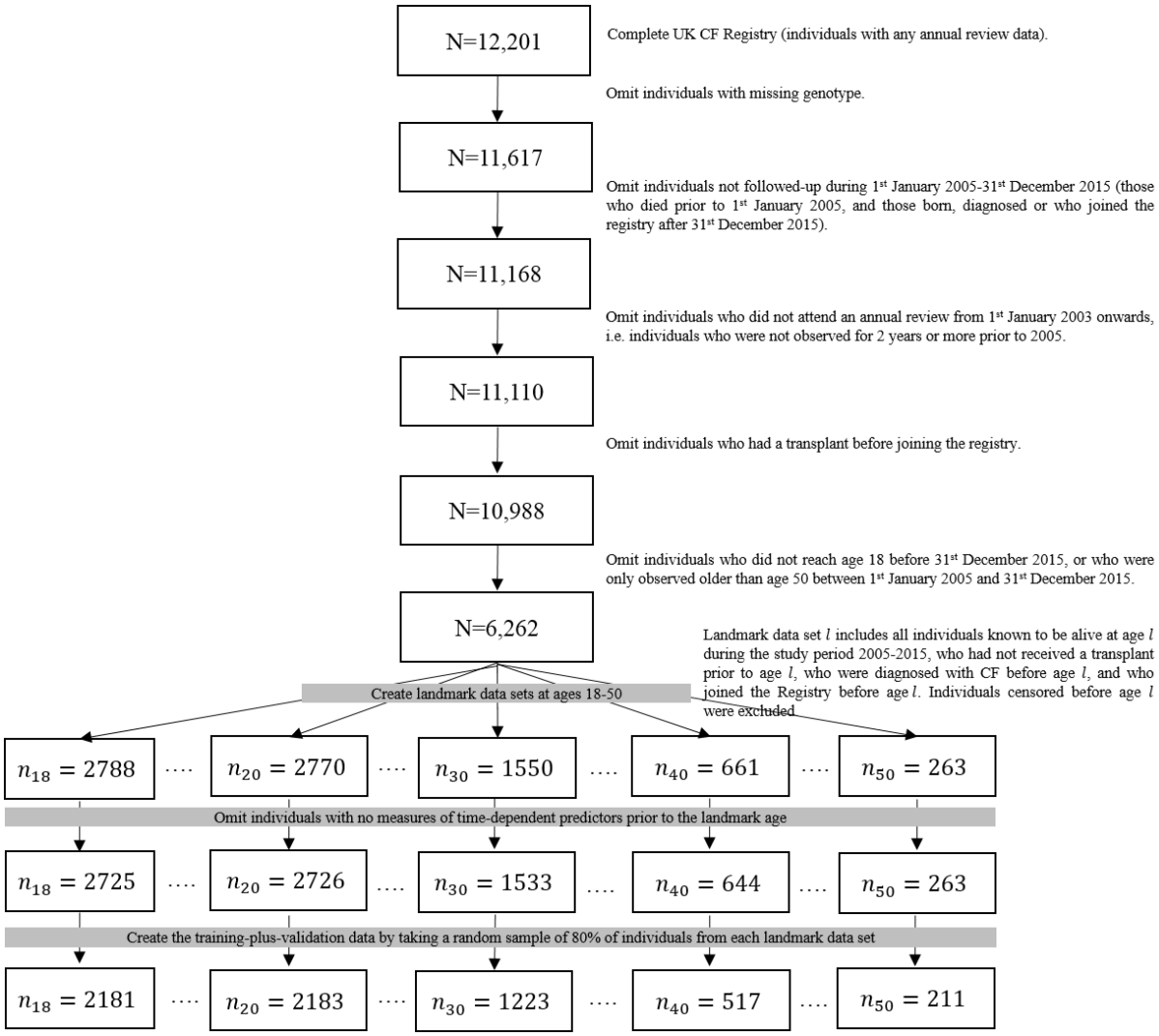
S6. Comparisons with other models

In an analysis of the French CF Registry Nkam et al reported a cross-validated C-statistic of 0.90 for prediction of 3-year survival.[16] They did not report a Brier score. Aside from focusing on 3-year survival and using different set of predictors, there are a number of differences between their approach and ours. They used a composite outcome of death and transplant, and for their logistic regression analysis, they excluded individuals who were censored before the end of the 3-year follow-up period.

Liou et al used a logistic regression analysis of the US CF Registry to predict 5-year survival.[17] A calibration plot showed good performance using a validation data set. However, they did not present measures of predictive performance that are comparable to those in this paper. Mayer-Hamblett et al also used a logistic regression analysis of the US CF Registry to develop a model for predicting 2-year survival.[18] They presented an ROC curve but did not report an area under the ROC curve, which could be compared to our C-Index. They presented sensitivities and specificities, and positive- and negative predictive values, finding that their model was better at predicting who would survive 2 years than who would die.

McCarthy et al developed the CF-ABLE score using logistic regression modelling of data from the CF population in Ireland.[19] Based on a validation data set, the area under the ROC curve was 0.82 for 4-year survival, though it is not clear how censoring was treated.

Supplementary Figure 1. Summary of data exclusions and creation of data set for analysis.



Supplementary Table 1. Summary of number of individuals, deaths, censorings and total person time at risk in each landmark data set. The stacked data set is formed by combining the landmark data sets.

Landmark age	No. of individuals	Number of deaths: N (%)			Number of censorings: N (%)		
		Within 2 years	Within 5 years	Within 10 years	Within 2 years	Within 5 years	Within 10 years
18	2725	63 (2.3)	171 (6.3)	255 (9.4)	500 (18.3)	1243 (45.6)	2290 (84.0)
19	2756	86 (3.1)	206 (7.5)	290 (10.5)	522 (18.9)	1244 (45.1)	2294 (83.2)
20	2726	104 (3.8)	218 (8.0)	303 (11.1)	505 (18.5)	1215 (44.6)	2239 (82.1)
21	2622	96 (3.7)	209 (8.0)	291 (11.1)	497 (19.0)	1221 (46.6)	2185 (83.3)
22	2526	107 (4.2)	206 (8.2)	273 (10.8)	477 (18.9)	1194 (47.3)	2104 (83.3)
23	2431	99 (4.1)	196 (8.1)	258 (10.6)	463 (19.0)	1159 (47.7)	2022 (83.2)
24	2326	85 (3.7)	182 (7.8)	234 (10.1)	501 (21.5)	1136 (48.8)	1970 (84.7)
25	2225	80 (3.6)	167 (7.5)	219 (9.8)	486 (21.8)	1088 (48.9)	1878 (84.4)
26	2079	82 (3.9)	160 (7.7)	216 (10.4)	439 (21.1)	1026 (49.4)	1760 (84.7)
27	1953	81 (4.1)	153 (7.8)	205 (10.5)	412 (21.1)	960 (49.2)	1647 (84.3)
28	1801	74 (4.1)	145 (8.1)	189 (10.5)	386 (21.4)	909 (50.5)	1540 (85.5)
29	1675	59 (3.5)	117 (7.0)	164 (9.8)	385 (23.0)	882 (52.7)	1436 (85.7)
30	1533	61 (4.0)	112 (7.3)	149 (9.7)	355 (23.2)	822 (53.6)	1323 (86.3)
31	1396	52 (3.7)	102 (7.3)	135 (9.7)	330 (23.6)	772 (55.3)	1205 (86.3)
32	1286	49 (3.8)	110 (8.6)	132 (10.3)	338 (26.3)	721 (56.1)	1112 (86.5)
33	1185	44 (3.7)	99 (8.4)	124 (10.5)	316 (26.7)	671 (56.6)	1011 (85.3)
34	1062	46 (4.3)	92 (8.7)	114 (10.7)	283 (26.6)	588 (55.4)	899 (84.7)
35	981	45 (4.6)	84 (8.6)	104 (10.6)	253 (25.8)	533 (54.3)	807 (82.3)
36	881	43 (4.9)	74 (8.4)	94 (10.7)	228 (25.9)	473 (53.7)	750 (85.1)
37	796	32 (4.0)	60 (7.5)	83 (10.4)	200 (25.1)	425 (53.4)	685 (86.1)
38	732	31 (4.2)	56 (7.7)	74 (10.1)	181 (24.7)	373 (51.0)	623 (85.1)
39	688	27 (3.9)	56 (8.1)	72 (10.5)	163 (23.7)	346 (50.3)	581 (84.4)
40	644	19 (3.0)	47 (7.3)	68 (10.6)	141 (21.9)	319 (49.5)	544 (84.5)
41	618	20 (3.2)	48 (7.8)	71 (11.5)	124 (20.1)	327 (52.9)	518 (83.8)
42	606	30 (5.0)	50 (8.3)	72 (11.9)	131 (21.6)	314 (51.8)	501 (82.7)
43	579	24 (4.1)	57 (9.8)	72 (12.4)	130 (22.5)	302 (52.2)	485 (83.8)
44	530	19 (3.6)	45 (8.5)	64 (12.1)	141 (26.6)	277 (52.3)	447 (84.3)
45	497	20 (4.0)	47 (9.5)	65 (13.1)	131 (26.4)	274 (55.1)	415 (83.5)
46	425	23 (5.4)	42 (9.9)	57 (13.4)	96 (22.6)	229 (53.9)	353 (83.1)
47	391	23 (5.9)	42 (10.7)	54 (13.8)	93 (23.8)	215 (55.0)	327 (83.6)
48	347	14 (4.0)	35 (10.1)	45 (13.0)	98 (28.2)	202 (58.2)	292 (84.1)
49	307	15 (4.9)	34 (11.1)	39 (12.7)	92 (30.0)	184 (59.9)	260 (84.7)
50	263	17 (6.5)	31 (11.8)	37 (14.1)	71 (27.0)	154 (58.6)	218 (82.9)

Supplementary Table 2. Summary of number of measurements of FEV1%, FVC% and weight used in multivariate mixed models fitted up to each landmark age. Results shown are the median, interquartile range (IAQR) and range of the number of measurements of each variable up to age l for individuals in the lth landmark data set ($l = 18, \ldots, 50$).

	FEV1%			FVC1%			Weight		
Landmark age	Median	IQR	Range	Median	IQR	Range	Median	IQR	Range
18	7	(5,10)	(1,20)	7	(5,10)	(1,20)	8	(5,11)	(1,21)
19	7	(5,10)	(1,21)	7	(5,10)	(1,21)	8	(5,11)	(1,22)
20	7	(5,10)	(1,22)	7	(5,10)	(1,22)	8	(5,11)	(1,22)
21	7	(5,10)	(1,21)	7	(5,10)	(1,21)	8	(5,11)	(1,21)
22	8	(5,10)	(1,21)	8	(5,10)	(1,21)	8	(5,11)	(1,21)
23	8	(5,10)	(1,22)	8	(5,10)	(1,22)	8	(5,11)	(1,22)
24	8	(5,11)	(1,24)	8	(5,11)	(1,24)	8	(6,11)	(1,24)
25	8	(6,11)	(1,24)	8	(5,11)	(1,24)	8	(6,11)	(1,24)
26	8	(5,11)	(1,24)	8	(5,11)	(1,24)	8	(6,11)	(1,25)
27	8	(5,11)	(1,23)	8	(5,11)	(1,23)	8	(6,11)	(1,22)
28	8	(6,11)	(1,23)	8	(6,11)	(1,23)	8	(6,11)	(1,23)
29	8	(5,11)	(1,22)	8	(5,11)	(1,22)	8	(6,11)	(1,22)
30	8	(6,11)	(1,20)	8	(6,11)	(1,20)	9	(6,11)	(1,21)
31	9	(6,11)	(1,21)	9	(6,11)	(1,21)	9	(6,12)	(1,21)
32	9	(6,11.75)	(1,23)	9	(6,11)	(1,23)	9	(6,12)	(1,23)
33	8	(5,11)	(1,21)	8	(5,11)	(1,21)	9	(5,12)	(1,22)
34	8	(5,11)	(1,23)	8	(5,11)	(1,23)	8	(5,12)	(1,24)
35	8	(5,11)	(1,19)	8	(5,11)	(1,19)	8	(5,12)	(1,19)
36	8	(5,11)	(1,19)	8	(5,11)	(1,19)	8	(5,12)	(1,19)
37	8	(5,11)	(1,19)	8	(5,11)	(1,19)	8	(5,11)	(1,18)
38	8	(5,11)	(1,19)	8	(5,11)	(1,18)	8	(5,11)	(1,18)
39	7	(4,11)	(1,18)	7	(4,11)	(1,18)	8	(4,11)	(1,19)
40	7	(4,10)	(1,19)	7	(4,10)	(1,19)	7	(4.5,11)	(1,18)
41	7	(4,10)	(1,18)	7	(4,10)	(1,18)	7	(5,11)	(1,18)
42	7	(4,10)	(1,17)	7	(4,10)	(1,17)	7	(4,10)	(1,17)
43	7	(4,10)	(1,18)	7	(4,10)	(1,18)	7	(4,10)	(1,18)
44	7	(5,10)	(1,19)	7	(5,10)	(1,19)	7	(5,11)	(1,19)
45	7	(5,10)	(1,20)	7	(5,10)	(1,20)	7	(5,10)	(1,20)
46	7	(5,10)	(1,18)	7	(5,10)	(1,18)	7	(5,10)	(1,18)
47	7	(5,10)	(1,19)	7	(5,10)	(1,19)	8	(5,10)	(1,18)
48	7	(5,10)	(1,20)	7	(5,10)	(1,20)	8	(5,10.5)	(1,20)
49	8	(5,11)	(1,21)	8	(5,11)	(1,21)	8	(5,11)	(1,20)
50	8	(5,11)	(1,16)	8	(5,11)	(1,16)	8	(5,11)	(1,16)

Supplementary Table 3. Results from the holdout data. Comparison between predicted survival probabilities from the final model and numbers of survivors and deaths within 2, 5 and 10 years from landmark ages 20, 30, 40 and 50. For 2-, 5-, and 10-year survival we excluded those who were censored before 2, 5 and 10 years of follow-up respectively. Note that due to small numbers in some predicted probability groups we do not expect the observed percentages surviving to exactly match the predicted survival probabilities.

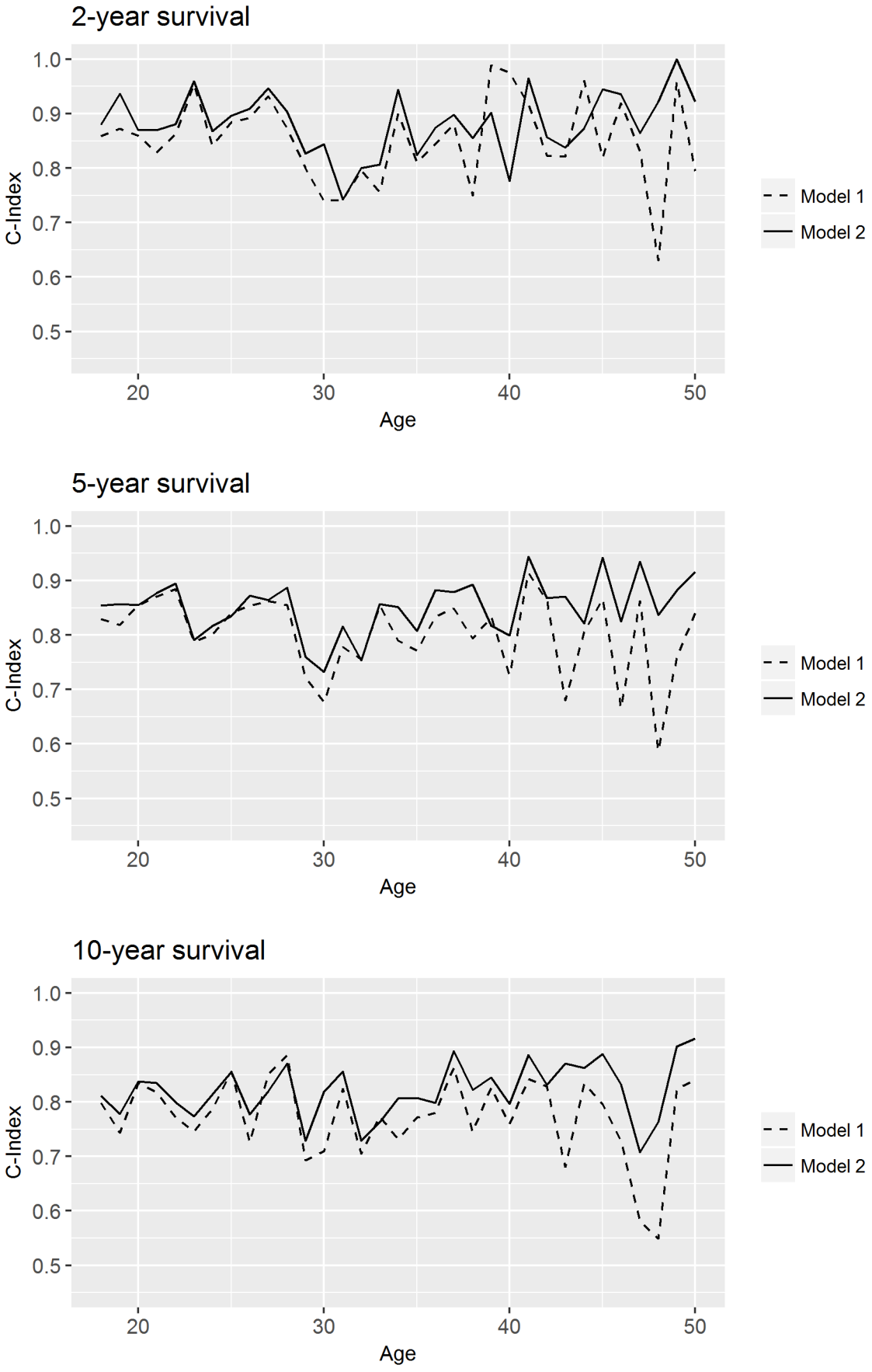
		2-year survival		5-year survival		10-year survival	
Landmark age	Probability of 2-year, 5-year or 10-year survival from final model	No. (%) who survived 2 years	No. (%) who died within 2 years	No. (%). who survived 5 years	No. (%) who died within 5 years	No. (%) who survived 10 years	No. (%) who died within 10 years
20	[0,0.7]	4 (57%)	3 (43%)	13 (48%)	14 (52%)	8 (20%)	32 (80%)
	(0.7,0.9]	43 (90%)	5 (10%)	51 (78%)	14 (22%)	13 (50%)	13 (50%)
	(0.9,0.95]	43 (93%)	3 (7%)	50 (94%)	3 (6%)	8 (73%)	3 (27%)
	(0.95,1]	341 (99%)	3 (1%)	166 (98%)	3 (2%)	15 (83%)	3 (17%)
30	[0,0.7]	2 (50%)	2 (50%)	5 (25%)	15 (75%)	3 (10%)	26 (90%)
	(0.7,0.9]	22 (73%)	8 (27%)	27 (73%)	10 (27%)	5 (28%)	13 (72%)
	(0.9,0.95]	37 (95%)	2 (5%)	21 (84%)	4 (16%)	4 (57%)	3 (43%)
	(0.95,1]	160 (99%)	2 (1%)	63 (97%)	2 (3%)	3 (75%)	1 (25%)
40	[0,0.7]	0 (0%)	1 (100%)	0 (0%)	5 (100%)	2 (14%)	12 (86%)
	(0.7,0.9]	5 (71%)	2 (29%)	17 (74%)	6 (26%)	1 (50%)	1 (50%)
	(0.9,0.95]	11 (100%)	0 (0%)	10 (100%)	0 (0%)	1 (100%)	0 (0%)
	(0.95,1]	81 (99%)	1 (1%)	31 (100%)	0 (0%)	0	0
50	[0,0.7]	3 (100%)	0 (0%)	1 (25%)	3 (75%)	0 (0%)	5 (100%)
	(0.7,0.9]	5 (71%)	2 (29%)	4 (67%)	2 (33%)	0 (0%)	2 (100%)
	(0.9,0.95]	8 (89%)	1 (11%)	5 (100%)	0 (0%)	1 (100%)	0 (0%)
	(0.95,1]	20 (100%)	0 (0%)	4 (100%)	0 (0%)	1 (100%)	0 (0%)
	[0,0.7]						

Supplementary Table 4. Overall C-Indexes and Brier scores for prediction of 2-year, 5-year and 10-year survival from a model including FEV1% as the only predictor and from a model including two treatment variables in addition to the 16 predictors included in the final model (Model 2 in Table 4 of the main text).

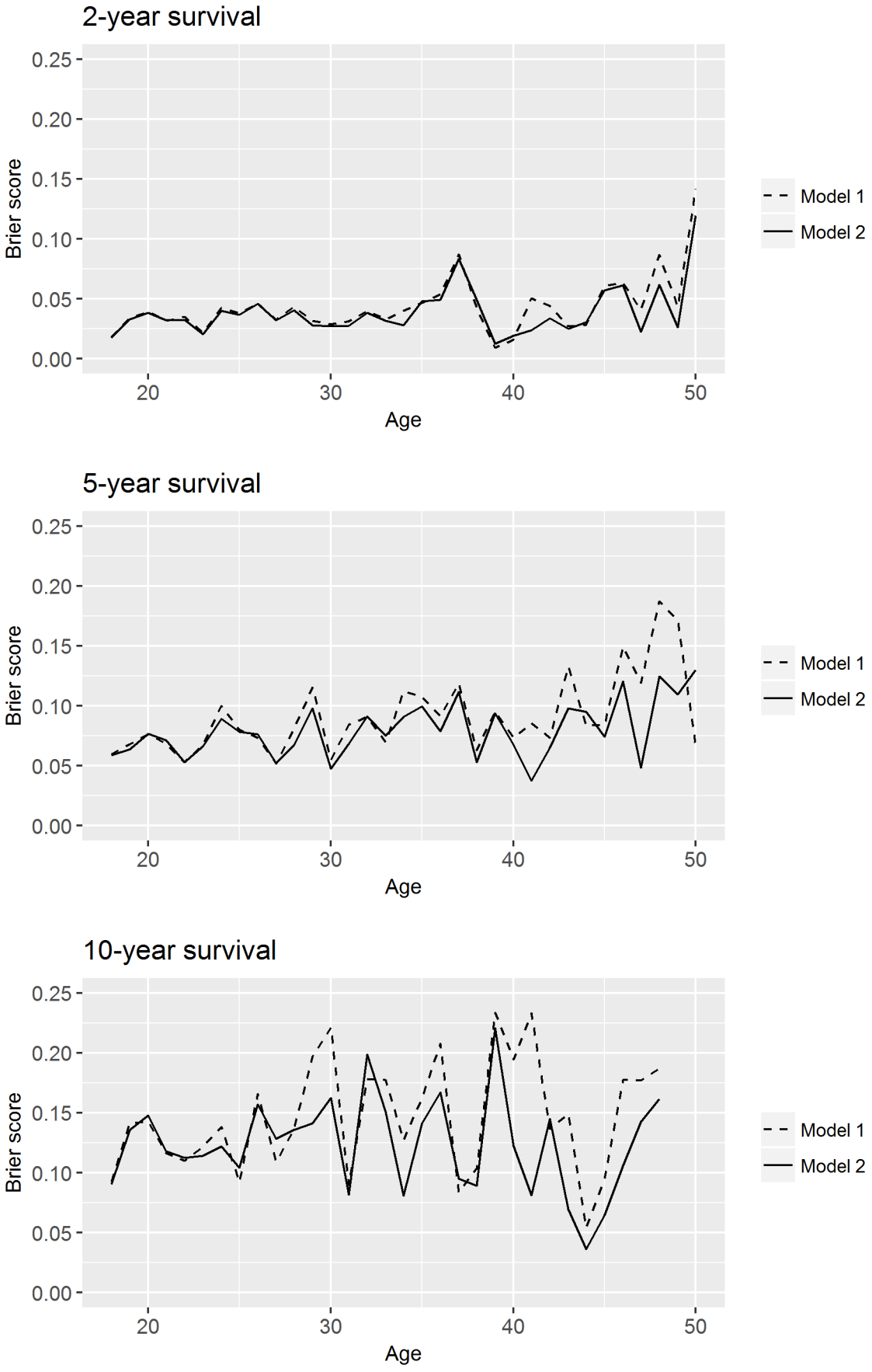
	Results from the final model (Model 2: Table 4 of the main text)		Model using FEV1% predicted as the only predictor[a]		Additionally including two treatment variables in Model 2[b]	
	C-Index	Brier score	C-Index	Brier score	C-Index	Brier score
2-year survival	0.873	0.036	0.842	0.038	0.876	0.035
5-year survival	0.843	0.076	0.813	0.081	0.844	0.075
10-year survival	0.804	0.133	0.775	0.141	0.805	0.133

a We repeated the final model with FEV1% predicted as the only predictor. Other features of the model were as in Model 2.
b We assessed the impact on predictive performance of including two treatments that were included in the model of Nkam et al for the French Registry: use of oxygen therapy and use of non-invasive ventilation.[16] Nkam et al also investigated use of oral corticosteroids, but there was insufficient data on use of this treatment in the UK data. We created binary variables at each landmark age, which indicate whether an individual had ever used each treatment in the past. The adjusted hazard ratio associated with oxygen use was 1.75 (95% CI 1.50-2.05) and the adjusted hazard ratio associated with non-invasive ventilation is 1.15 (95% CI 0.92-1.43). Therefore both oxygen therapy and non-invasive ventilation are associated with an increased mortality hazard (though the association for non-invasive ventilation is not statistically significant), because these treatments are used by sicker patients. The estimates do not have a causal interpretation.

Supplementary Figure 2. Comparison of landmark-age-specific C-indexes for 2-year, 5-year and 10-year survival from Model 1 (separate models from each landmark age) and Model 2 (supermodel).



Supplementary Figure 3. Comparison of landmark-age-specific Brier scores for 2-year, 5-year and 10-year survival from Model 1 (separate models from each landmark age) and Model 2 (supermodel).



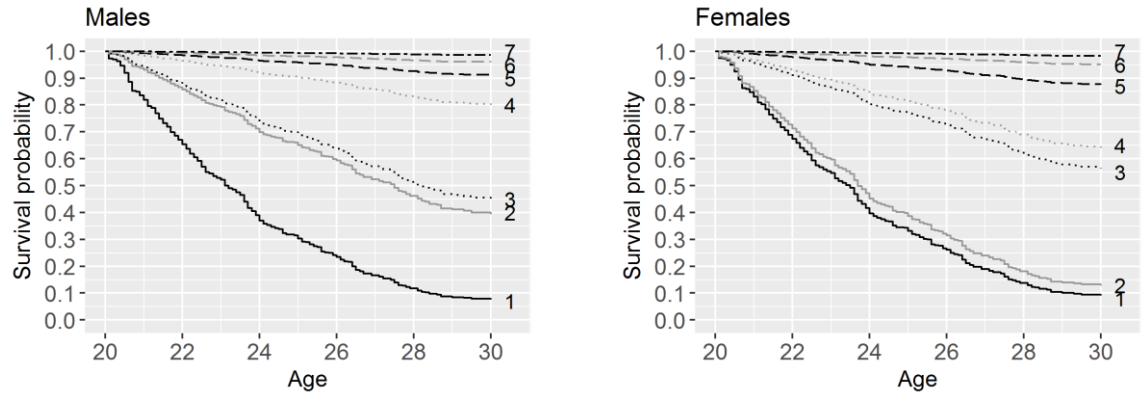
Supplementary Figure 4. Predicted survival curves from landmark age 20 for example individuals in groups defined by 5-year survival probabilities. For individuals in the Registry at age 20 between 2013 and 2015 we obtained their predicted 5-year survival probabilities and categorized into groups with 5-year survival probabilities <0.5, (0.5,0.7,], (0.7,0.8], (0.8,0.9], (0.9,0.95], (0.95,0.99], (0.99, 1]. An example individual was created for each group.

(i) Characteristics of example individuals[a] in groups defined by 5-year survival probability.

5-year survival probability group	<0.5	(0.5,0.7]	(0.7,0.8]	(0.8,0.9]	(0.9,0.95]	(0.95,0.99]	(0.99,1]
Example person	1	2	3	4	5	6	7
Males, Females[b]							
Genotype (no. copies of F508del)	1, 2	2	2	2	2	2	1
Age of diagnosis (years)	0	0	0	0	0	0	1, 3
FEV1%	22, 26	28, 35	43, 39	44, 51	59, 61	78, 70	97, 98
FVC%	32, 39	48. 53	58, 54	64, 72	75, 77	90, 90	103, 106
Weight (kg)	48, 46	53, 47	51, 49	56, 48	57, 53	65, 56	73, 64
Height (cm)	167, 159	169, 156	167, 160	174, 158	170, 158	173, 161	177, 164
P. aeruginosa	Yes	Yes	Yes	Yes	Yes	Yes	No
B. cepacia	No	No	No	No	No	No	No
S. aureus	No	Yes, No	Yes, No	No	Yes, No	No	No
MRSA	No	No	No	No	No	No	No
Pancreatic insufficiency	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CF related diabetes	Yes	Yes	Yes	No, Yes	No	No	No
Hospitalisation (not for IVs)	No	No	No	No	No	No	No
Number of hospital IV days	29+	15-28, 29+	29+, 1-14	1-14, 15-28	0	0	0
Number of hospital IV days	0	1-14, 29+	0	0, 1-14	0	0	0

[a]We created an example individual for each group using the median values of the continuous predictors and the most common value of each categorical variable within that group. For hospital and home IV days we obtained the median number of days and then assigned the relevant category. This was done separately for males and females.
[b] Values are shown as 'male, female', except were the value for males and females was the same.

(ii) Predicted survivor curves based on the final model for example individuals with characteristics shown in the table above.



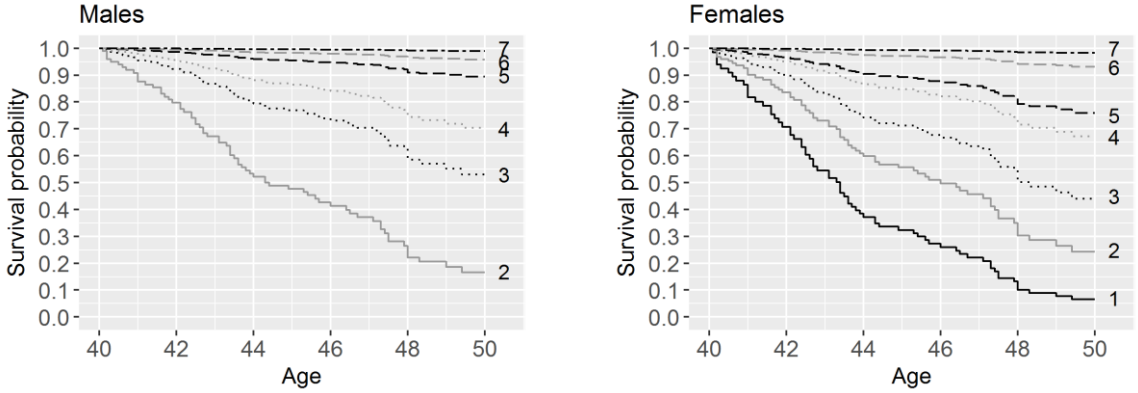
Supplementary Figure 5. Predicted survival curves from landmark age 40 for example individuals in groups defined by 5-year survival probabilities. For individuals in the Registry at age 40 between 2013 and 2015 we obtained their predicted 5-year survival probabilities and categorized into groups with 5-year survival probabilities <0.5, (0.5,0.7,], (0.7,0.8], (0.8,0.9], (0.9,0.95], (0.95,0.99], (0.99, 1]. An example individual was created for each group.

(i) Characteristics of example individuals[a] in groups defined by 5-year survival probability. Results are not shown for groups of less than 5 individuals.

5-year survival probability group	<0.5	(0.5,0.7]	(0.7,0.8]	(0.8,0.9]	(0.9,0.95]	(0.95,0.99]	(0.99,1]
Example person	1	2	3	4	5	6	7
Males, Females[b]							
Genotype (no. copies of F508del)	-	2	2	2,1	2,1	1	1
Age of diagnosis (years)	-	2, 0	1, 0	1, 4	3, 3	2, 14	29, 13
FEV1%	-	27, 25	31, 28	38, 41	51, 47	68, 65	92, 92
FVC%	-	42, 43	60, 45	64, 59	70, 66	93, 81	97, 96
Weight (kg)	-	67	64	63	68	75	85
Height (cm)	-	173	170	173	176	176	175
P. aeruginosa	-	Yes	Yes	Yes	Yes	Yes	No
B. cepacia	-	No	No	No	No	No	No
S. aureus	-	No	No	No	No	No	No, Yes
MRSA	-	No	No	No	No	No	No
Pancreatic insufficiency	-	Yes	Yes	Yes	Yes	Yes	No
CF related diabetes	-	Yes	Yes	Yes	No, Yes	No	No
Hospitalisation (not for IVs)	-	No	No	No	No	No	No
Number of hospital IV days	-	29+	15-28, 1-14	1-14	0, 1-14	0	0
Number of hospital IV days	-	29+, 1-14	0, 1-14	0	0	0	0

[a]We created an example individual for each group using the median values of the continuous predictors and the most common value of each categorical variable within that group. For hospital and home IV days we obtained the median number of days and then assigned the relevant category. This was done separately for males and females.
[b] Values are shown as 'male, female', except were the value for males and females was the same.

(ii) Predicted survivor curves based on the final model for example individuals with characteristics shown in the table above.



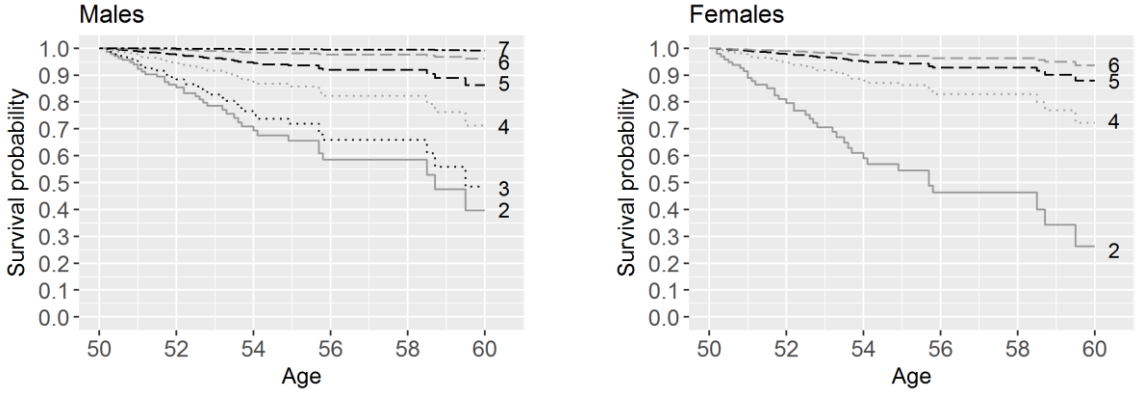
Supplementary Figure 6. Predicted survival curves from landmark age 50 for example individuals in groups defined by 5-year survival probabilities. For individuals in the Registry at age 50 between 2013 and 2015 we obtained their predicted 5-year survival probabilities and categorized into groups with 5-year survival probabilities <0.5, (0.5,0.7,], (0.7,0.8], (0.8,0.9], (0.9,0.95], (0.95,0.99], (0.99, 1]. An example individual was created for each group.

(i) Characteristics of example individuals[a] in groups defined by 5-year survival probability. Results are not shown for groups of less than 5 individuals.

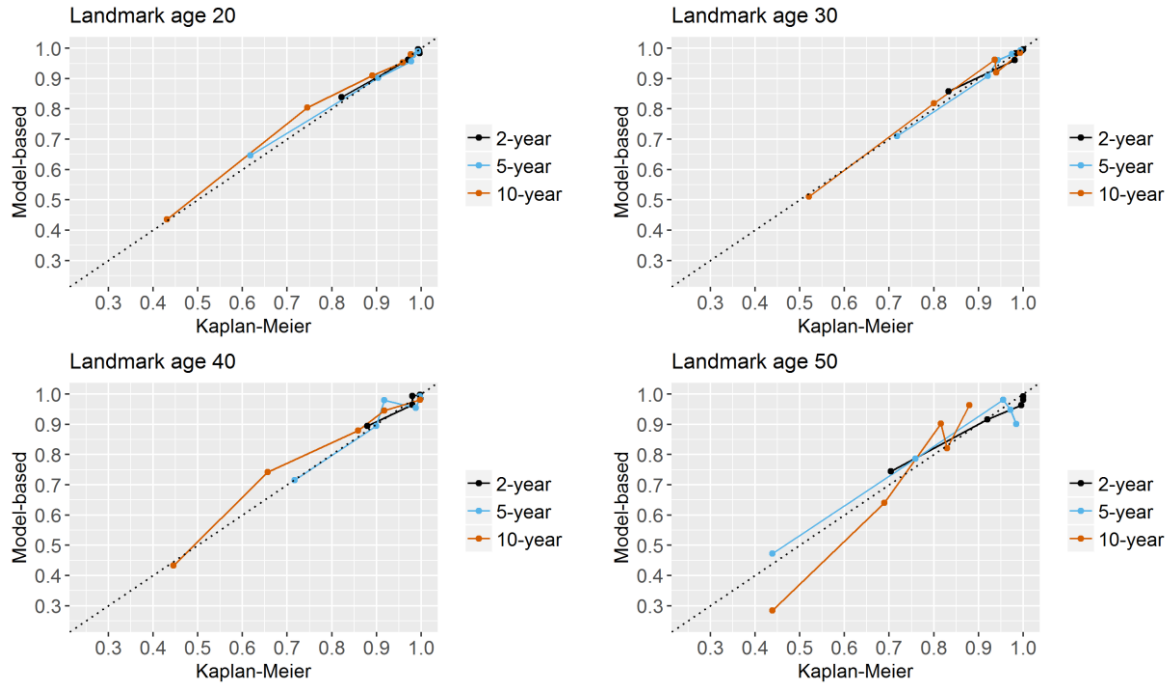
5-year survival probability group	<0.5	(0.5,0.7]	(0.7,0.8]	(0.8,0.9]	(0.9,0.95]	(0.95,0.99]	(0.99,1]
Example person	1	2	3	4	5	6	7
Males/Females[b]							
Genotype (no. copies of F508del)	-	1, 2	2, -	2, 1	2, 1	1	1, -
Age of diagnosis (years)	-	4, 1	4, -	1	6, 28	28, 34	39, -
FEV1%	-	31, 30	27, -	48, 49	55, 64	82, 76	113, -
FVC%	-	51, 64	63, -	72, 69	76, 81	91, 90	108, -
Weight (kg)	-	65, 55	76, -	76, 61	80, 66	79, 65	86, -
Height (cm)	-	172, 158	174, -	17, 165	176, 162	176, 163	177, -
P. aeruginosa	-	Yes	Yes, -	Yes	No, Yes	No	No, -
B. cepacia	-	No	No, -	No	No	No	No, -
S. aureus	-	No	No, -	No	No	No	No, -
MRSA	-	No	No, -	No	No	No	No, -
Pancreatic insufficiency	-	Yes	Yes, -	Yes	Yes	Yes, No	No, -
CF related diabetes	-	Yes	Yes, -	Yes, No	No	No	No, -
Hospitalisation (not for IVs)	-	No	No, -	No	No	No	No, -
Number of hospital IV days	-	1-14	1-14, -	0	0	0	0, -
Number of hospital IV days	-	0	1-14, -	0	1-14	0	0, -

[a]We created an example individual for each group using the median values of the continuous predictors and the most common value of each categorical variable within that group. For hospital and home IV days we obtained the median number of days and then assigned the relevant category. This was done separately for males and females.
[b] Values are shown as 'male, female', except were the value for males and females was the same.

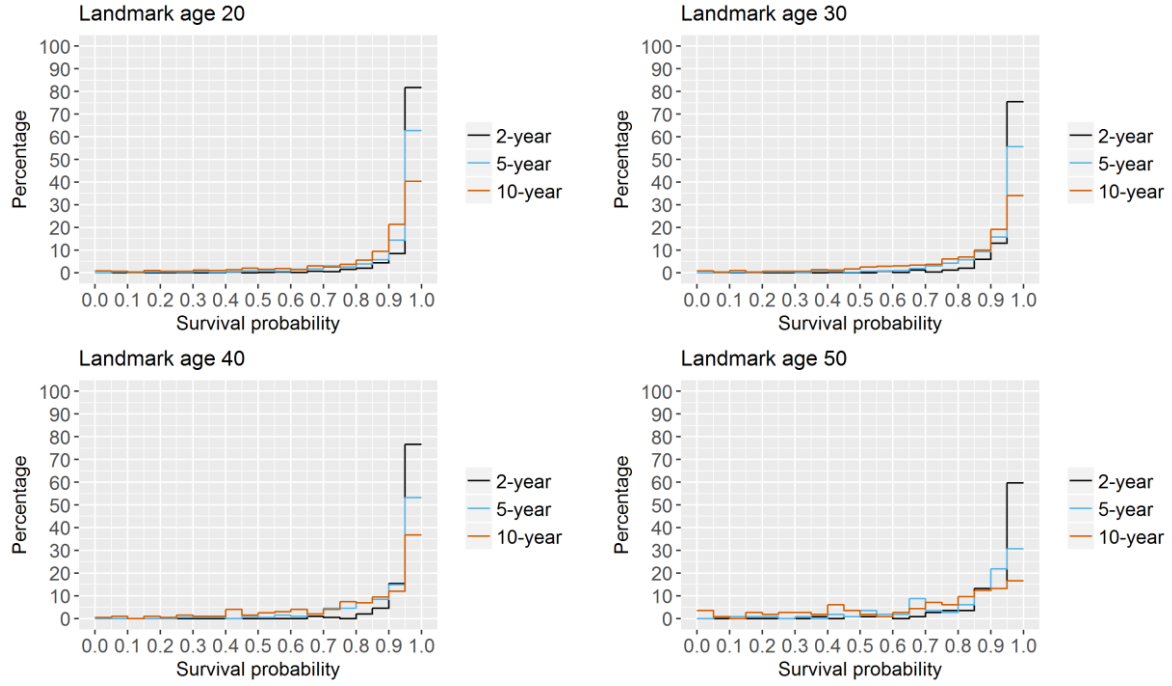
(ii) Predicted survivor curves based on the final model for example individuals with characteristics shown in the table above.



Supplementary Figure 7. [This is a colour version of Figure 2 in the main text.] Calibration plots using the final model (Model 2) for prediction of 2-year, 5-year and 10-year survival from landmark ages 20, 30, 40 and 50. The vertical axis shows the mean model-based x-year survival probability (x=2,5,10) in quintiles of the model-based probabilities. The horizontal axis shows the mean x-year survival probability obtained using Kaplan-Meier estimates in quintiles of the model-based probabilities. The five points have been joined by a line.



Supplementary Figure 8. [This is a colour version of Figure 4 in the main text.] 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.



References

1. Cox, D. R. Models and Life-Tables Regression. *J. R. Stat. Soc. Ser. B.* **34,** 187–220 (1972).
2. Cox, D. Partial likelihood. *Biometrika* **62,** 269–276 (1975).
3. Breslow, N. Discussion of the paper by D. R. Cox. *J. R. Stat. Soc. Ser. B* **34,** 216–217 (1972).
4. 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).
5. 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).
6. 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).
7. 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).
8. Antolini, L., Boracchi, P. & Biganzoli, E. A time-dependent discrimination index for survival data. *Stat. Med.* **24,** 3927–3944 (2005).
9. Graf, E., Schmoor, C., Sauerbrei, W. & Schumacher, M. Assessment and comparison of prognostic classification schemes for survival data. *Stat. Med.* **18,** 2529–2545 (1999).
10. 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).
11. van Houwelingen, H. C. & Putter, H. *Dynamic prediction in clinical survival analysis*. (CRC Press/Chapman and Hall, 2012).
12. Kuhn, M. & Johnson, K. Over-Fitting and Model Tuning. in *Applied Predictive Modelling* 61–92 (Springer, 2013).
13. Therneau, T. A Package for Survival Analysis in S. version 2.38. (2015).
14. Putter, H. Dyn pred package. (2015).
15. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R. C. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131.1. (2018).
16. Nkam, L. *et al.* A 3-year prognostic score for adults with cystic fibrosis. *J. Cyst. Fibros.* **119,** 682–689 (2017).
17. Liou, T. G. *et al.* Predictive 5-year survivorship model of cystic fibrosis. *Am. J. Epidemiol.* **153,** 345–352 (2001).
18. 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).
19. 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 cystic fibrosis. *Chest* **143,** 1358–1364 (2013).