

# Survival models to support shared decision-making about advance care planning for people with advanced stage cystic fibrosis

Negin Hajizadeh <sup>1,2,3</sup> Meng Zhang,<sup>3,4</sup> Meredith Akerman,<sup>4</sup> Nina Kohn,<sup>4</sup> Anna Mathew,<sup>5</sup> Denis Hadjiiladis,<sup>6</sup> Janice Wang,<sup>1,2</sup> Martin L Lesser<sup>3,4</sup>

**To cite:** Hajizadeh N, Zhang M, Akerman M, *et al*. Survival models to support shared decision-making about advance care planning for people with advanced stage cystic fibrosis. *BMJ Open Res Res* 2021;**8**:e000794. doi:10.1136/bmjresp-2020-000794

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjresp-2020-000794>).

Received 5 October 2020  
Revised 18 March 2021  
Accepted 2 May 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Negin Hajizadeh;  
nhajizadeh@northwell.edu

## ABSTRACT

**Background** For people with advanced stage cystic fibrosis (CF), tailored survival estimates could facilitate preparation for decision-making in the event of acutely deteriorating respiratory function.

**Methods** We used the US CF Foundation national database (2008–2013) to identify adult people with incident advanced stage CF (forced expiratory volume in 1 s (FEV1)  $\leq$ 45% predicted). Using the lasso method for variable selection, we divided the dataset into training and validation samples (2:1), and developed two multivariable Cox proportional hazards models to calculate probabilities of survival from baseline (T0 model), and from 1 year after (T12 model). We also performed Kaplan-Meier survival analyses.

**Results** 4752 people were included. For the T0 model, FEV1; insurance; non-invasive ventilation; supplemental oxygen; *Burkholderia* colonisation; cirrhosis; depression; dialysis; current smoking; unclassifiable mutation class and cumulative CF exacerbations predicted increased mortality. Baseline transplant evaluation status of ‘accepted, on waiting list’ predicted decreased mortality. For the T12 model, interim decrease in FEV1  $>$ 10%, and pulmonary exacerbations additionally increased predicted mortality. Lung transplantation was associated with lower mortality. Of the 4752, 93.5%, 86.4%, 79.7% and 73.9% survived to 1, 2, 3 and 4 years, respectively, without considering any confounding variables. The models had moderate predictive ability indicated by the area under the time-dependent receiver operating characteristic curve (0.787, 95% CI 0.769 to 0.794 for T0 model; and 0.779, 95% CI 0.767 to 0.797 for T12 model).

**Conclusion** We have developed models predicting survival in people with incident advanced stage CF, which can be reapplied over time to support shared decision-making about end-of-life treatment choices and lung transplantation. These estimates must be updated as data become available regarding long-term outcomes for people treated with CF transmembrane conductance regulator modulators.

## INTRODUCTION

People with cystic fibrosis (CF) and their families feel unprepared to discuss life supporting treatment decisions at the time when people

## Key messages

### What is the key question?

► What are the predictors of 12, 24 and 36 months mortality for people with incident advanced stage cystic fibrosis (CF) and how do these predictors change 12 months after incident advanced stage CF?

### What is the bottom line?

► Lung function (exacerbation rates, forced expiratory volume in 1 s, bacterial colonisation) and other end-organ dysfunction predicts mortality for people with incident advanced stage CF; at 1-year follow-up, decline in lung function predicts increased mortality, and lung transplantation was associated with decreased mortality.

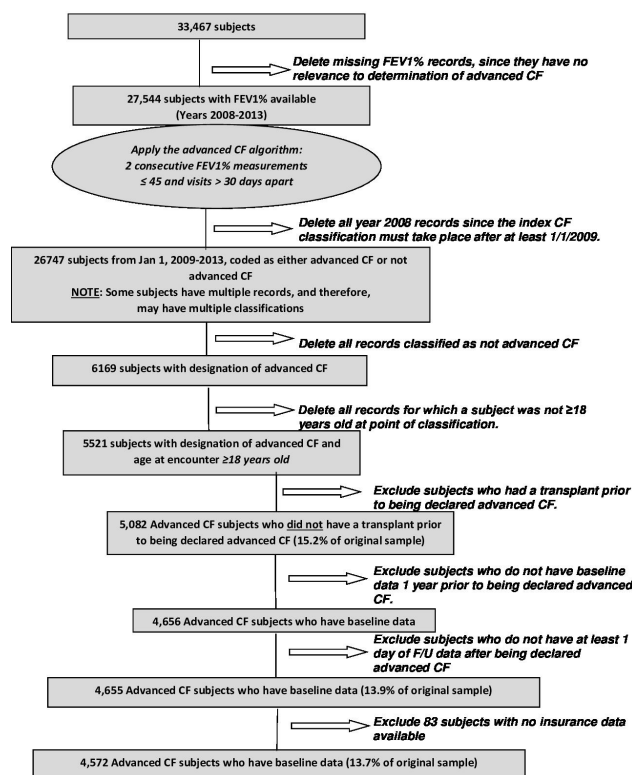
### Why read on?

► We have developed practical survival models that can be applied during clinic visits to inform iterative conversations with updated survival predictions over time for people with advanced stage CF.

suffer acute respiratory failure. People with CF with severe lung function impairment (advanced stage CF) are at particular risk of respiratory failure. Long-term outcomes for those who are not able to receive lung transplantation are poor and often result in prolonged end-of-life care in the intensive care unit (ICU).<sup>1–3</sup> Many patients receive invasive life supporting technologies without the opportunity to weigh the risks and benefits of such therapies, as these decisions often need to be made in the midst of an emergency, and in a heightened state of fear and anxiety. Very often, the first mention of mechanical ventilation occurs in the hospital with providers who have no established rapport with the patient and who may not have an understanding of the unique needs and mindset of people with CF.

Predicting survival for people with advanced stage CF remains difficult. Prognostic ability is limited by the relatively high degree of interpatient variability and the natural course of pulmonary decline.<sup>4–9</sup> CF prediction tools are hampered by complexity and difficulty in practical day-to-day usage.<sup>7,9</sup> As such, decisions are guided by individual clinical judgement, and conversations in advance of critical illness are often delayed. An easy-to-use index that could accurately provide tailored mortality estimates in people with advanced stage CF would be helpful to both patients and clinicians, particularly if integrated within shared decision aids to support advance care planning.<sup>10–12</sup> McCarthy *et al* have developed the CF-ABLE12 which uses clinical parameters that are measured at every clinic visit.<sup>13</sup> However, it was developed in Ireland where healthcare provision and cultural differences might make this index less applicable to the US-based CF population. Liou *et al* derived and validated a 5-year survivorship model using patient registry data from the Cystic Fibrosis Foundation (CFF) in the USA.<sup>14</sup> However, this model is now outdated since the patient data used was from 1986 to 1997 and the model is difficult to use in practice. Critical care and CF care have both improved in recent years, and continue to improve. For example, non-invasive ventilation is now routinely used in advanced CF for exacerbations, and extracorporeal membrane oxygenation is now offered as a bridge to transplantation. Most recently, Aaron *et al* published the results of a survival model in 2015 based on the Canadian CF patient registry, and Nkam *et al* published a model based on the French CF registry (externally validated using Canadian registry data).<sup>15–17</sup> In the USA, MacKenzie *et al* published updated survival trends using US CF patient registry data from 2000 to 2010.<sup>18</sup> While these existing models aid in prognostic discussions, they are not specific to advanced stages of lung disease. In addition, with advances in CF treatments, and medical care in general, prognostic models to inform conversations about treatment choices need to be continually updated. Furthermore, models must consider individual countries' healthcare systems and populations.

Based on these considerations, our group sought to develop a survival model which could inform shared decision-making conversations between clinicians and their patients with advanced stage CF who are at higher risk for respiratory failure and mortality. These patients could benefit most from preparation to decide about life supporting technologies including invasive mechanical ventilation and lung transplantation. The model was developed for use with patients whose lung function has deteriorated to FEV1  $\leq$ 45% predicted, and for iterative use at yearly follow-up visits. It has been incorporated into a shared decision aid that was co-developed with patient with CF, caregiver and clinician stakeholder input.



**Figure 1** Study flow diagram. CF, cystic fibrosis; FEV1, forced expiratory volume in 1 s.

## METHODS

### Patient and public involvement

Patients were not involved in this study.

We analysed data from 1 January 2008 to 31 December 2013 from the CFF, which maintains a registry on all people with CF in the USA ( $n=33\,467$  subjects). In order to have at least 1 year of baseline data, we limited the data to people with incident advanced stage CF in 2009 or later (defined as two consecutive events of FEV1  $\leq$ 45% predicted at least 30 days apart) and age  $>18$  years. We chose the FEV1 cut-off based on the definition used by the CFF for lung disease severity,<sup>19</sup> and the cut-off point for FEV1 that was associated with higher mortality in the CF-ABLE study.<sup>13</sup> We excluded people who had a lung transplant prior to incident advanced stage CF, and people who did not have baseline data 1 year prior to becoming advanced stage CF. Further details of cohort selection are outlined in our Consolidated Standards of Reporting Trials diagram (figure 1).

Models were built to inform a baseline conversation between clinicians and patients when patients were identified as having lung function decline suggesting advanced stage CF (FEV1  $\leq$ 45%), and again 12 months later at a follow-up visit. We aimed to provide clinicians, and their patients, tailored estimates for 12, 24 and 36 months survival probability. To this end, we created two separate multivariable Cox proportional hazards models<sup>1</sup>: The 'T0' model—a baseline model to calculate the survival probability (which we report for 12, 24 and 36 months of follow-up), and<sup>2</sup> The 'T12' model—a

model to calculate the conditional probability of additional survival for people who survived 12 months after becoming advanced stage (again reported for 12, 24 and 36 months of follow-up). Survival times were considered censored if a subject was alive as of the last day available in the database (31 December 2013).

### Variable selection

Selection of candidate variables for the multivariable model proceeded as follows: first, from the available data we selected 30 potential predictor variables based on prior literature and clinical expertise.<sup>16–20</sup> Next, for categorical variables with K categories, K-1 dummy variables were included in the model development. Each variable was standardised by subtracting its mean and dividing by its SD. For the training data, the lasso method, adapted to survival data was then applied to all the variables using a 10-fold cross-validation approach.<sup>21</sup> Lasso is a modern method of selecting candidate variables which prevents overfitting of a model which, in turn, can result in poor performance of prediction. This yielded the optimal value of the tuning parameter,  $\lambda^*$ , such that the mean cross-validated error was within 1 SD of the minimum error. Using  $\lambda^*$  we then performed lasso variable selection on the training data, which yielded estimated beta-coefficients for the multivariable model. The variable selection procedure was carried out using the R package ‘glmnet’ and ‘coxnet’.<sup>22–23</sup>

It should be noted that the beta-coefficient estimated for a given predictor variable using lasso is ‘biased’, which means that it does not represent the true magnitude of effect. In other words, unlike in cox regression models that use classical techniques of variable selection, exponentiation of the lasso beta-coefficient is not an estimate of the true HR. However, when taken in context of the model, it can determine relative importance in prediction and positive or negative association with the predicted outcome. Although the lasso model gives biased estimates, it is superior to the classical techniques because the classic techniques may not get the absolute best combination of predictors when the predictors are correlated, which may affect the accuracy of prediction. In comparison, lasso takes all the predictor variables into account through regularisation which is more efficient, so that the accuracy of prediction is increased.

The ‘T0’ and ‘T12’ samples (n=4572 and n=3822, respectively) were each randomly divided into a training and validation sample in a 2:1 fashion. At time 0, the training sample size was  $n_0=3047$  and validation sample size was  $n_0=1525$ . At time 12, the training sample size was  $n_{12}=2548$  and validation sample size was  $n_{12}=1274$ . The respective training samples were used to derive the T0 and T12 models, and the validation samples were used to validate the predictive ability of the respective models.

### Modelling survival at T0

The T0 model estimates the probability of 12, 24 and 36 months survival from when the patient is identified as having advanced stage CF, for a given combination of covariates. Continuous variables with non-monotone trend in the Kaplan-Meier survival curve were transformed into categorical variables. Only those variables with non-zero estimated beta-coefficients were included in the final model.

### Modelling survival at T12

The T12 model estimates the probability of (an additional) 12, 24 and 36 months survival from after a patient has survived 12 months for a given combination of covariates. Selection of variables for the T12 model was carried out using the same methods as the model for T0.

### Validation of the T0 and T12 models

In order to obtain an unbiased assessment of the goodness-of-fit of the T0 and T12 models, the derived models were applied to the validation samples, and we computed the area under the time-dependent receiver operating characteristic curve (AUC) (function of survival ROC.C in R package survivalROC).<sup>24</sup> AUC is used to validate the predictive ability of a survival model; the larger the AUC the better the discriminatory ability, where 0 denotes no predictive capacity at all and 1 denotes perfect prediction.

### Estimation of 12, 24 and 36 months survival probabilities

Although the Cox proportional hazards models were developed to estimate the survival probability to any time point, because of widening CIs beyond 36 months we chose to focus on 12, 24 and 36 months predictions. The estimated survival probabilities were computed using R package ‘glmnet’ and ‘coxnet’ and ‘Tibshirani’s and Allison’s Method’.<sup>22–23</sup> Details are in Online supplemental appendix 1 and hypothetical scenario results are in tables 3A and B.

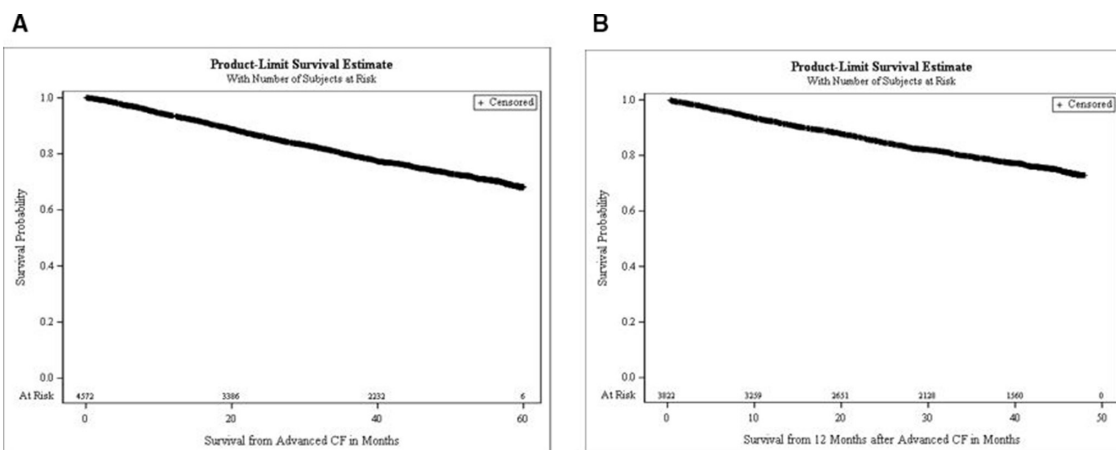
### Kaplan-Meier survival analysis

We performed Kaplan-Meier survival analyses in order to obtain crude estimates of survival without considering any confounding variables for the overall cohorts. These were performed for the T0 and T12 cohorts as well as for those who received transplant (figures 2 and 3)

## RESULTS

There were 4572 people who met the criteria for incident advanced stage CF in the period 1 January 2009 to 31 December 2013. Of this baseline cohort, 3822 people survived to 1 year. Demographic and clinical characteristics for the cohorts used in the T0 (baseline) and T12 (12 months) models are given in tables 1A and B. For the T0 model, the mean FEV1 was 34.2% predicted, mean age was 30.5 years and mean body mass index (BMI) was 21.0, and





**Figure 2** Kaplan-Meier survival curves for T0 cohort (A) and T1 cohort (B). CF, cystic fibrosis.

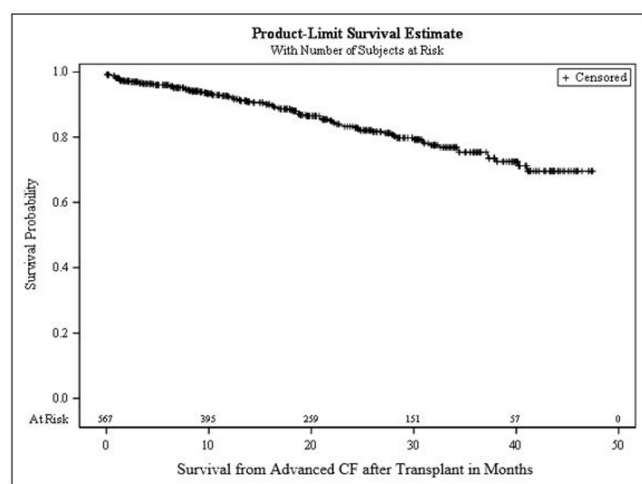
47.6% were female. Of the cohort, 31.4% were prescribed supplemental oxygen therapy, 12.9% were prescribed oral steroids, 53.1% had supplemental feeding, 72.2% had mutation class 1–3 and 22.4% were recorded as disabled. Almost all people were white (94.9%).

### Cox proportional hazards models

The estimated beta-coefficients of the multivariable Cox proportional hazards models T0 and T12 using the lasso method are shown in tables 2A and B and examples of combinations of covariates for hypothetical people are presented in tables 3A and B.

Results of regularised (lasso) Cox proportional hazards model. Variables predictive of mortality for T0 model (table 2A) and T12 model (table 2B).

Estimated beta-coefficients represent the results of the multivariable Cox proportional hazards model predicting mortality over the course of the available study follow-up period. Only variables with non-zero beta-coefficients are included.



**Figure 3** Kaplan-Meier survival curve for patients who received lung transplantation. CF, cystic fibrosis.

For the T0 model, the following variables had non-zero beta-coefficients and were associated with increased risk of mortality: decreased baseline FEV1 (% predicted); having a combination of insurance including Medicaid compared with having private insurance; non-invasive ventilation; continuous oxygen therapy; *Burkholderia* species in sputum culture; cirrhosis; depression; renal failure requiring dialysis; current smoking; unclassifiable mutation class at baseline ('other') and cumulative CF exacerbations in the year prior to incident advanced stage CF. Having baseline transplant evaluation status of 'accepted, on waiting list' predicted decreased mortality as compared with 'not pertinent'.

For the T12 model (1 year after incident advanced stage CF), the following variables continued to be associated with increased predicted mortality: having a combination of insurance including Medicaid compared with having private insurance; continuous oxygen therapy; depression and having increased cumulative CF exacerbations in the year prior to incident advanced stage CF. Additionally, increased mortality was predicted for those with interim deterioration in FEV1 over the 1 year of follow-up by >10%, and for those with higher number of interim pulmonary exacerbations. Lung transplantation over the 1 year of follow-up was associated with lower mortality compared with no transplantation.

**Validation:** when the fitted models were applied to the validation samples, area under the time-dependent ROC curve was 0.787 (95% CI 0.769 to 0.794) for the T0 model and 0.779 (95% CI 0.767 to 0.797) for the T12 model, which indicates a moderate level of predictive ability. The AUC of the validation sample for the model with FEV1 alone was 0.668 (95% CI 0.653 to 0.682) for the T0 model and 0.657 (95% CI 0.641 to 0.672) for T12 model.

### Estimation of 12, 24 and 36 months survival probabilities

In order to give examples of how these models could be applied in clinical scenarios, we created three hypothetical patient scenarios, using combinations of covariables, for which we assumed, based on clinical experience and

**Table 1A** Patient characteristics at baseline for T0 model

	Training (n=3047)	Validation (n=1525)	Overall sample (n=4572)
FEV1 (% predicted)	34.2 (7.6)	34.3 (7.3)	34.2 (7.6)
Gender (female)	1420 (46.6%)	758 (49.7%)	2178 (47.6%)
Age at incident advance CF	30.5±10.7 (median=28.0)	30.4±11.1 (median=27.0)	30.5±10.8 (median=27.0)
Race (white)	2885 (94.7%)	1456 (95.5%)	4341 (94.9%)
Insurance			
Private insurance only	1340 (44.0%)	690 (45.2%)	2030 (44.4%)
Medicare only	125 (4.1%)	76 (5.0%)	201 (4.4%)
Combination of any insurance including Medicaid	1582 (51.9%)	759 (49.8%)	2341 (51.2%)
BMI	21.0±3.5 (median=20.4)	21.0±3.5 (median=20.3)	21.0±3.5 (median=20.3)
Hispanic ethnicity	148 (4.9%)	79 (5.2%)	227 (5.0%)
Ventilation machine (non-invasive)			
No	2770 (93.8%)	1391 (93.9%)	4161 (93.9%)
Yes	114 (3.9%)	57 (3.8%)	171 (3.9%)
Unknown	68 (2.3%)	33 (2.2%)	101 (2.3%)
Corticosteroids (oral (eg, prednisone))	392 (12.9%)	200 (13.1%)	592 (12.9%)
Patient is on enzymes	2719 (89.2%)	1362 (89.3%)	4081 (89.3%)
Oxygen therapy			
No	1964 (66.5%)	970 (65.5%)	2934 (66.2%)
Yes, continuously	130 (4.4%)	61 (4.1%)	191 (4.3%)
Yes, nocturnal and/or with exertion	265 (9.0%)	138 (9.3%)	403 (9.1%)
Yes, during exacerbation	451 (15.2%)	239 (16.1%)	690 (15.6%)
Yes, prn	95 (3.2%)	50 (3.4%)	145 (3.3%)
Unknown	47 (1.6%)	23 (1.6%)	70 (1.6%)
Supplemental feeding	1601 (52.5%)	825 (54.1%)	2426 (53.1%)
<i>Haemophilus influenzae</i> (any species)	197 (6.5%)	85 (5.6%)	282 (6.2%)
<i>Pseudomonas aeruginosa</i> (found in the culture)	2305 (75.7%)	1173 (76.9%)	3478 (76.1%)
Burkholderia complex ( <i>Burkholderia</i> species)			
No	2795 (94.7%)	1407 (95.0%)	4202 (94.8%)
Yes	157 (5.3%)	74 (5.0%)	231 (5.2%)
Other microorganisms	650 (21.3%)	325 (21.3%)	975 (21.3%)
<i>Staphylococcus aureus</i> /resistant to MRSA:*			
+/+	861 (28.3%)	426 (27.9%)	1287 (28.1%)
+/-	793 (26.0%)	391 (25.6%)	1184 (25.9%)
-/NA	1304 (42.8%)	667 (43.7%)	1971 (43.1%)
<i>P. aeruginosa</i> /Resistant to aminoglycosides:†			
+/+	819 (26.9%)	413 (27.1%)	1232 (26.9%)
+/-	1486 (48.8%)	760 (49.8%)	2246 (49.1%)
-/NA	653 (21.4%)	311 (20.39%)	964 (21.1%)
Pulmonary complications (massive haemoptysis or pneumothorax requiring chest tube)	126 (4.1%)	68 (4.5%)	194 (4.2%)
Liver disease, cirrhosis			
No	2851 (96.6%)	1434 (96.8%)	4285 (96.7%)
Yes	101 (3.4%)	47 (3.2%)	148 (3.3%)

Continued

Table 1A Continued

	Training (n=3047)	Validation (n=1525)	Overall sample (n=4572)
Cancer confirmed by histology	16 (0.5%)	6 (0.4%)	22 (0.5%)
Depression			
No	2178 (73.8%)	1118 (75.5%)	3296 (74.4%)
Yes	774 (26.2%)	363 (24.5%)	1137 (25.6%)
Renal failure requiring dialysis			
No	2939 (99.6%)	1474 (99.5%)	4413 (99.5%)
Yes	13 (0.4%)	7 (0.5%)	20 (0.5%)
# of clinic visits in the review year	4.4±3.0 (median=4.0)	4.5±2.9 (median=4.0)	4.5±2.9 (median=4.0)
Disabled	668 (21.9%)	357 (23.4%)	1025 (22.4%)
DM or impaired glucose tolerance	1351 (44.3%)	681 (44.7%)	2032 (44.4%)
Smoking			
No	2865 (97.1%)	1435 (96.9%)	4300 (97.0%)
Occasionally	41 (1.4%)	20 (1.4%)	61 (1.4%)
Regularly, <1 ppd	27 (0.9%)	23 (1.6%)	23 (0.5%)
Regularly, 1 ppd or more	14 (0.5%)	1 (0.07%)	1 (0.02%)
Unknown	5 (0.2%)	2 (0.1%)	2 (0.05%)
Mutation class			
1–3	2207 (72.4%)	1095 (71.8%)	3302 (72.2%)
4–5	218 (7.2%)	114 (7.5%)	332 (7.3%)
Other	429 (14.1%)	222 (14.6%)	651 (14.2%)
Unknown missing	193 (6.3%)	94 (6.2%)	287 (6.3%)
# of pulmonary exacerbations in the year preceding advanced stage of diagnosis	2.0±2.0 (median=2.0)	2.0±2.0 (median=2.0)	2.0±2.0 (median=2.0)
Baseline lung transplant evaluation status:			
Not pertinent	2808 (92.2%)	1407 (92.3%)	4215 (92.2%)
Accepted, on waiting list	43 (1.4%)	17 (1.1%)	60 (1.3%)
Evaluated, final decision pending	80 (2.6%)	42 (2.8%)	122 (2.7%)
Evaluated, rejected	11 (0.3%)	12 (0.8%)	23 (0.5%)
Unknown	10 (0.3%)	3 (0.2%)	13 (0.3%)

\*Refers to presence of MRSA on culture/resistant to vancomycin.

†Refers to presence of *Pseudomonas* on culture/resistant to all aminoglycoside.

BMI, body mass index; CF, cystic fibrosis; DM, diabetes mellitus; FEV1, forced expiratory volume in 1 s; MRSA, methicillin-resistant *S. aureus*; NA, not available.

Table 1B Patient characteristics at baseline for T12 model

	Training (n=2548)	Validation (n=1274)	Overall sample (n=3822)
FEV1	34.4 (7.3)	34.3 (7.4)	34.4 (7.3)
Interim change in FEV1 (T12–T0)*			
Greater than >10% (improvement in FEV1)	782 (31.6%)	416 (33.8%)	1198 (32.3%)
Between –10% and 10%	1083 (43.8%)	514 (41.8%)	1597 (43.1%)
Less than –10% (deterioration in FEV1)	608 (24.6%)	300 (24.4%)	908 (24.5%)
Gender (female)	1209 (47.4%)	594 (46.6%)	1803 (47.2%)
Age at incident advance CF	30.8±10.7 (median=28.0)	30.7±10.9 (median=28.0)	30.7±10.8 (median=28.0)

Continued

Table 1B Continued

	Training (n=2548)	Validation (n=1274)	Overall sample (n=3822)
Race (white)	2419 (94.9%)	1213 (95.2%)	3632 (95.0%)
Insurance			
Private insurance only	1174 (46.1%)	551 (43.2%)	1725 (45.1%)
Medicare only	111 (4.4%)	59 (4.6%)	170 (4.4%)
Combination of any insurance including Medicaid	1263 (49.6%)	664 (52.1%)	1927 (50.4%)
BMI	21.1±3.6 (median=20.4)	21.0±3.5 (median=20.3)	21.1±3.6 (median=20.4)
Hispanic ethnicity	114 (4.5%)	66 (5.2%)	180 (4.7%)
Ventilation machine (non-invasive)	75 (2.9%)	42 (3.3%)	117 (3.1%)
Corticosteroids (oral (eg, prednisone))	311 (12.2%)	160 (12.6%)	471 (12.3%)
Patient is on enzymes	2275 (89.3%)	1128 (88.5%)	3403 (89.0%)
Oxygen therapy			
No	1707 (69.0%)	811 (65.9%)	2518 (68.0%)
Yes, continuously	90 (3.6%)	43 (3.5%)	133 (3.6%)
Yes, nocturnal and/or with exertion	211 (8.5%)	119 (9.7%)	330 (8.9%)
Yes, during exacerbation	339 (13.7%)	204 (16.6%)	543 (14.7%)
Yes, prn	86 (3.5%)	37 (3.0%)	123 (3.3%)
Unknown	40 (1.6%)	16 (1.3%)	56 (1.6%)
Supplemental feeding	1306 (51.3%)	675 (53.0%)	1981 (51.8%)
<i>Haemophilus influenzae</i> (any species)	174 (6.8%)	73 (5.7%)	247 (6.5%)
<i>Pseudomonas aeruginosa</i> (found in the culture)	1930 (75.7%)	981 (77.0%)	2911 (76.2%)
Burkho complex ( <i>Burkholderia</i> species)	125 (4.9%)	58 (4.5%)	183 (4.8%)
Other microorganisms	522 (20.5%)	289 (22.7%)	811 (21.2%)
<i>Staphylococcus aureus</i> /MRSA†			
+/+	697 (27.3%)	341 (26.7%)	1038 (27.2%)
+/-	672 (26.4%)	320 (25.1%)	992 (25.9%)
-/NA	1108 (43.5%)	571 (44.8%)	1679 (43.9%)
<i>Pseudomonas aeruginosa</i> /Resistant to aminoglycosides‡			
+/+	690 (27.1%)	345 (27.1%)	1035 (27.1%)
+/-	1240 (48.7%)	636 (49.9%)	1876 (49.1%)
-/NA	547 (21.5%)	251 (19.7%)	798 (20.9%)
Pulmonary complications (massive haemoptysis or pneumothorax requiring chest tube)	115 (4.5%)	49 (3.8%)	164 (4.3%)
Liver disease, cirrhosis	85 (3.3%)	28 (2.2%)	113 (3.0%)
Cancer confirmed by histology	13 (0.5%)	4 (0.3%)	17 (0.4%)
Depression			
No	1843 (74.5%)	943 (76.7%)	2786 (75.2%)
Yes	630 (25.5%)	287 (23.3%)	917 (24.8%)
Renal failure requiring dialysis	13 (0.5%)	4 (0.3%)	17 (0.4%)
# of clinic visits in the review year	4.4±2.9 (median=4.0)	4.4±2.9 (median=4.0)	4.4±2.9 (median=4.0)
Disabled§	571 (22.4%)	268 (21.0%)	839 (21.9%)
DM or impaired glucose tolerance	1062 (41.7%)	579 (45.4%)	1641 (42.9%)
Smoking	65 (2.5%)	35 (2.7%)	100 (2.6%)
Mutation class			

Continued



Table 1B Continued

	Training (n=2548)	Validation (n=1274)	Overall sample (n=3822)
1–3	1850 (72.6%)	921 (72.3%)	2771 (72.5%)
4–5	197 (7.7%)	89 (7.0%)	286 (7.5%)
Other	349 (13.7%)	189 (14.8%)	538 (14.1%)
Unknown missing	152 (6.0%)	75 (5.9%)	277 (5.9%)
# of pulmonary exacerbations in the year preceding advanced stage CF diagnosis	1.8±1.9 (median=1.0)	2.1±2.1 (median=2.0)	1.9±1.9 (median=1.0)
Interim number of pulmonary exacerbations (T0–T12)	1.8±1.8 (median=1.0)	1.9±2.0 (median=1.0)	1.8±1.9 (median=1.0)
Lung transplant in the past 12 months			
No	2123 (85.8%)	1037 (84.3%)	3160 (85.3%)
Yes	350 (14.2%)	193 (15.7%)	543 (14.7%)

\*FEV1% change calculated as  $([FEV1\%T12 - FEV1\%T0] / [FEV1\%T0]) \times 100$ .

†Refers to presence of MRSA on culture/resistant to vancomycin.

‡Refers to presence of *Pseudomonas* on culture/resistant to all aminoglycoside.

§Refers to patient unemployed due to a disability.

BMI, body mass index; CF, cystic fibrosis; DM, diabetes mellitus; FEV1, forced expiratory volume in 1 s; MRSA, methicillin-resistant *S. aureus*; NA, not available.

representation in the database, there would be high, moderate and low survival probabilities. Combinations with lower FEV1, combined insurance, continuous oxygen therapy, liver disease and cirrhosis, missing mutation class, increased numbers of pulmonary exacerbations and baseline lung transplant evaluation status of 'not pertinent' had lower survival probability for the T0 model. Combinations with worsening of FEV1 between baseline and 12 months, combined insurance, depression, increased numbers of pulmonary exacerbations, having baseline lung transplant evaluation status as 'not pertinent' and having no lung transplant in the past 12 months had lower survival probability for T12 model, as well as a steeper trajectory of decline in survival over time. However, the CIs for the scenarios with lower predicted survival probability were wide, likely explained by the number of people with these combinations in the cohort, or by the presence of potential outliers.

### Descriptive results of people listed for transplant

Of the 4572 people in the study cohort, 939 (21%) were listed on the lung transplant waiting list at some time during the 4-year study period, either at the time they became advanced CF (n=60) or subsequent to that date. During the study follow-up time period, 567 (60%) of these people received a lung transplant, on average within 1 year of being listed for transplant. Of the 567 people who received a lung transplant, 7% received a transplant during the 1 year after becoming advanced stage CF; 27% between years 1 and 2; 29% between years 2 and 3; 22% between years 3 and 4 and 15% between years 4 and 5. Of the 372 who were listed but did not receive a transplant by the end of the study follow-up time, 171 (15%) were alive on the waiting list, 120 (13%) were listed as 'other'

(including unknown or missing, not pertinent, final decision pending or rejected), and 81 (9%) died while on the waiting list.

### Kaplan-Meier survival plots

Survival plot results for the T0 and T12 cohorts are presented in figure 2. Figure 3 depicts the survival plot for those who received a lung transplant. Of the 4572 people with incident advanced stage CF, 93.6%, 86.4%, 79.8% and 73.9% survived to 1, 2, 3 and 4 years, respectively without considering any confounding variables. Of the 3822 people who survived at least 12 months after becoming incident advanced stage CF, 92.3%, 85.3%, 79.0% and 72.9% survived an additional 1, 2, 3 and 4 years, respectively without considering any confounding variables. Of the 567 people who received a lung transplant, 92.2%, 83.2%, 75.4% survived to 1, 2 and 3 years after transplant, respectively, without considering any confounding variables.

### DISCUSSION

We have developed survival models for people with CF with incident advanced stage lung disease with the goal of prompting conversations about planning for invasive mechanical ventilation and lung transplantation. Our study was designed to create usable survival models which can calculate survival probabilities over time and be updated at 1-year follow-up visits, specifically for people with incident advanced stage CF who are in greater need of advance care planning.

Our model found similar results to prior models estimating survival in people with CF. As expected, lower lung function, smoking, *Burkholderia cepacia* colonisation,



**Table 2A** Variables predictive of mortality for T0 model

	Beta-coefficient from lasso*
FEV1 (% predicted)	-0.044
Insurance	
Private insurance only	Reference group
Medicare only	0
Combination of any insurance including Medicaid	0.16
Ventilation machine (non-invasive)	
No	Reference group
Yes	0.103
Unknown	0
Oxygen therapy	
No	Reference group
Yes, continuously	0.46
Yes, nocturnal and/or with exertion	0
Yes, during exacerbation	0
Yes, prn	0
Unknown	0
Burkho complex ( <i>Burkholderia</i> species)	
No	Reference group
Yes	0.315
Liver disease, cirrhosis	
No	Reference group
Yes	0.116
Renal failure requiring dialysis	
No	Reference group
Yes	0.584
Smoking	
No	Reference group
Occasionally	0.141
Regularly, <1 ppd	0.429
Regularly, 1 ppd or more	0
Unknown	0
Mutation class	
1-3	Reference group
4-5	0
Other	0
Unknown missing	0.302

Continued

**Table 2A** Continued

	Beta-coefficient from lasso*
# of pulmonary exacerbations in the year preceding advanced stage CF diagnosis	0.093
Baseline lung transplant evaluation status	
Not pertinent	Reference group
Accepted, on waiting list	-0.102
Evaluated, final decision pending	0
Evaluated, rejected	0
Unknown	0

Parameter estimates do not represent the true magnitude of effect. In other words, unlike in Cox regression models that use classical techniques of variable selection, exponentiation of the lasso beta-coefficient is not an estimate of the true HR. However, when taken in context of the model it can determine relative importance in prediction and positive or negative association with the predicted outcome.

\*A negative value indicates predictive of *lower* mortality. CF, cystic fibrosis; FEV1, forced expiratory volume in 1 s.

**Table 2B** Variables predictive of mortality for T12 model

	Beta-coefficient from lasso
Insurance	
Private insurance only	Reference group
Medicare only	0
Combination of any insurance including Medicaid	0.051
Oxygen therapy	
No	Reference group
Yes, continuously	0.459
Yes, nocturnal and/or with exertion	0
Yes, during exacerbation	0
Yes, prn	0
Unknown	0
Depression	
No	Reference group
Yes	0.006
# of pulmonary exacerbations in the year preceding advanced stage CF diagnosis	0.016
Interim change in FEV1 (T12-T0)	
Improvement >10% (+10%)	Reference group
Between -10% and 10%	0
Deterioration >10% (-10%)	0.445
Interim number of pulmonary exacerbations (T0-T12)	0.131
Lung transplant in the past 12 months	

Continued



Table 2B Continued

	Beta-coefficient from lasso
No	Reference group
Yes	-0.494*

Parameter estimates do not represent the true magnitude of effect. In other words, unlike in Cox regression models that use classical techniques of variable selection, exponentiation of the lasso beta-coefficient is not an estimate of the true HR. However, when taken in context of the model it can determine relative importance in prediction and positive or negative association with the predicted outcome.

\*A negative value indicates predictive of lower mortality. CF, cystic fibrosis; FEV1, forced expiratory volume in 1 s.

multiple exacerbations and need for non-invasive ventilation were associated with worse survival. In addition, rapid lung function decline in the first year of advanced lung disease was associated with worse outcomes. Notably, our model showed worse survival for people with 'combination insurance'. This category includes those on Medicaid, representing lower socioeconomic status, which is a recognised risk factor for poor health outcomes. Depression, cirrhosis and renal failure requiring renal replacement therapies were the specific comorbidities found to predict increased mortality. However, unlike other models, BMI and age were not found to be associated with worse survival in the multivariable models. This is very likely due to the association with mortality captured by other variables included in the model, such as lung function decline, oxygen supplementation and non-invasive ventilation. The increased mortality seen with the 'other mutation class' is harder to explain. This class was present in 14% of the sample and represents those in whom the mutation class could not be defined at the time of entry.

As survival models are updated, it is important to compare findings with other models. The CF-ABLE score includes age, BMI, FEV1 and number of exacerbations, is based on a 49-person Irish CF cohort, with a maximum 6-year follow-up and validated with a 370-person Irish CF cohort. The mean FEV1 in both training and validation cohorts was 62% predicted (95% CI 55% to 69%), and 60% predicted (95% CI 58% to 63%), respectively. Of this cohort, 13 (27%) died or received a lung transplant which were combined as one outcome (3 deaths and 10 transplantations). The area under the ROC curve was 0.82 (95% CI 0.77 to 0.88). A score based on the French CF registry identified FEV1, BMI, *B. cepacia* colonisation, number of intravenous antibiotics, days of hospitalisation, the need for oral corticosteroids, long-term oxygen therapy and non-invasive ventilation as important predictors of mortality. This was based on a cohort of 2096 people with CF, with a maximum of 3 years of follow-up data, between 2010 and 2013. The mean FEV1 was 58.3%

predicted (95% CI 39.4% to 79.8%). Of this cohort, 268 (13%) died or received a lung transplant which were combined as one outcome (55 deaths and 213 transplantation). By comparison, the mean FEV1 was 34% predicted in our cohort, 24% died by the end of the study period and 12% received lung transplantation. We explored the association between lung transplantation and survival by considering lung transplant as a time-dependent covariate in our models, however, we found no significant association. This is likely due to the severity of illness associated with the receipt of a lung transplant being captured by other variables in the model, such as number of exacerbations, supplemental therapies needed, comorbidities, FEV1 predicted and change in FEV1 predicted over time.

We have validated the discriminative ability of our model using conventional splitting of the data into training and validation samples. Our c-statistic for this model demonstrates moderate discriminatory ability that it can discriminate between pairs of patients where one died (experienced the event of interest) and the other did not die, and predicts the patient with the lower risk score as being the one who did not die. These models are not intended to exactly predict the likelihood of survival but rather to communicate relative differences in risk between patients with different characteristics. One explanation for the moderate predictive ability of our model could be that the correct model was not employed. However, the Cox proportional hazards model is known to be a flexible model because it is semi-parametric and non-parametric, and we verified the proportional hazards assumptions. An alternative explanation is that other types of data are needed for better prediction, such as biomarkers or variability in vital signs. In addition, as seen in the hypothetical scenarios (tables 3A and B), some combinations of covariables are infrequently represented and therefore we do not have enough data to inform the model for those scenarios.

The main limitation of our model is due to the granularity of data available within the dataset from which the models are derived. As epigenetic and other modifiers of outcome become apparent, these variables must be integrated within prognostic models. A further limitation is the potentially incomplete data on people after they have received transplantation, as many people seek care with transplant clinics which do not necessarily input data to the CFF registry. Additionally, assumptions made for models are always limiting and we have attempted to transparently outline each of our assumptions for review. The FEV1 choice for 'advanced stage' CF was based on clinical expertise using a cut-off that is above that used typically for lung transplant referral. This was intentionally chosen to allow for time to discuss with loved ones, to review information carefully selected for educating people (including patient and caregiver narratives) within the decision aid. We realise that this cut-off is very slightly higher than the FEV1 <40% defined in the recent

**Table 3A** Examples of predicted probabilities of survival to 12, 24 and 36 months for the T0 cohort (ie, after baseline incident advanced CF): three different hypothetical examples of people with extreme combinations of covariables

Scenarios (covariables)	Predicted probability of survival to 12 months given covariables X (95% CI)	Predicted probability of survival to 24 months given covariables X (95% CI)	Predicted probability of survival to 36 months given covariables X (95% CI)
FEV1 15% predicted; Combination of insurance (Medicaid); Non-smoker; No non-invasive ventilation; Continuous oxygen therapy; No <i>Burkholderia</i> species No liver disease; Depression; No renal failure requiring dialysis; Mutation class: 1–3; 4 or more pulmonary exacerbations in the year preceding advanced stage CF diagnosis; Baseline lung transplant evaluation status: not pertinent (n=3)	77.9% (71.6% to 83.2%)	58.2% (48.3% to 65.7%)	43.5% (32.7% to 51.2%)
FEV1 25% predicted; Combination of insurance (Medicaid); Non-smoker; No non-invasive ventilation; Continuous oxygen therapy; No <i>Burkholderia</i> species; No liver disease; No depression; No renal failure requiring dialysis; Mutation class: 1–3; 4 or more pulmonary exacerbations in the year preceding advanced stage CF diagnosis; Baseline lung transplant evaluation status: not pertinent (n=21)	86.7% (83.4% to 90.1%)	73.4% (68.2% to 77.5%)	62.2% (55.6% to 66.9%)
FEV1 40 (%) predicted; Private insurance only; Non-smoker; No non-invasive ventilation; No oxygen therapy; No <i>Burkholderia</i> species; No liver disease No depression; No renal failure requiring dialysis; Mutation class: 1–3, 0 pulmonary exacerbations in the year preceding advanced stage CF diagnosis; Baseline lung transplant evaluation status: not pertinent (n=131)	97.3% (96.7% to 98.1%)	94.1% (92.9% to 95.8%)	90.8% (88.9% to 93.1%)

n=X represents the number of people with the combinations of covariables in the dataset.  
CF, cystic fibrosis; FEV1, forced expiratory volume in 1 s.

**Table 3B** Examples of predicted probabilities of survival to 12, 24 and 36 months for the T12 cohort (ie, for people surviving to 12 months after baseline): three different hypothetical examples of people with extreme combinations of covariables

Scenarios (covariables)	Predicted probability of survival to 12 months given covariables X (95% CI)	Predicted probability of survival to 24 months given covariables X (95% CI)	Predicted probability of survival to 36 months given covariables X (95% CI)
Deterioration in FEV1 (T12–T0) >10%; Combination of insurance (Medicaid); No oxygen therapy; Depression; 4 or more pulmonary exacerbations in the year preceding advanced stage CF diagnosis; 4 or more interim pulmonary exacerbations (T0–T12); No lung transplant in the past 12 months (n=7)	79.2% (69.3% to 84.1%)	64.1% (49.9% to 71.2%)	50.8% (33.8% to 59.4%)

Continued



Table 3B Continued

Scenarios (covariables)	Predicted probability of survival to 12 months given covariables X (95% CI)	Predicted probability of survival to 24 months given covariables X (95% CI)	Predicted probability of survival to 36 months given covariables X (95% CI)
Improvement in FEV1 (T12–T0) >10%; Combination of insurance (Medicaid); Continuous oxygen therapy; No depression; 4 or more pulmonary exacerbations in the year preceding advanced stage CF diagnosis 4 or more interim pulmonary exacerbations (T0–T12); No lung transplant in the past 12 months (n=25)	86.2% (82.4% to 90.9%)	75.3% (68.9% to 83.9%)	64.9% (57.7% to 75.5%)
Improvement in FEV1 (T12–T0) >10%; Private insurance only; No oxygen therapy; No depression No pulmonary exacerbations in the year preceding advanced stage CF diagnosis; No interim pulmonary exacerbations (T0–T12); No lung transplant in the past 12 months (n=128)	95.2% (94.5% to 97.9%)	91.0% (90.1% to 95.8%)	86.6% (84.7% to 93.4%)

n=X represents the number of people with the combinations of covariables in the dataset.  
CF, cystic fibrosis; FEV1, forced expiratory volume in 1 s.

Cystic Fibrosis Advanced Lung Disease guidelines,<sup>25</sup> which was published prior to the start of our analyses, but we do not believe this diminishes the value and clinical applicability of our model. Our model does not capture the increased survival resulting from the new modulating medications used for CF. As we describe in our introduction, prognostic models need to be iteratively updated to reflect changes in therapies over time with newer data when they become available as these novel therapeutics are integrated in to care over the next decade. In the meantime, it is possible that our model could be used to guide intensification and optimisation of therapy, and timely referral for lung transplant and initiation of goals of care and advanced care planning for people identified by the model to have high estimated mortality. Once data become available that include several years of outcomes for those receiving CF transmembrane conductance regulator (CFTR) modulators, the principles used to develop these models will be re-applied for updated models. In fact, this process of updating models should continue over time with progressive changes in treatment. The new CFTR modifiers have likely changed the prediction of survival significantly, however, the factors that lead to worse outcomes are likely to remain the same. Therefore, our models can be used to identify high-risk patients; their disease trajectory might have improved, but their early identification may still help clinicians and patients proactively engage in shared decision-making conversation about advance care planning and transplant referral. Finally, our models do not include prediction of quality of life outcomes which are important for decision-making about treatments, in addition to survival prediction. These outcomes are difficult to capture reliably from databases that are not designed to record quality of life metrics including patient-reported outcomes. However,

our models are intended to be applied within shared decision-making conversations, wherein clinicians can describe impact on quality of life. Furthermore, narratives within the decision aid which has been developed describe outcomes that impact quality of life—although without the statistical estimates of likelihood of such specific outcomes.

The strengths of our model include the potential to update estimates over time for people, and to depict the predicted survival over time for a given combination of covariates which are readily available during clinical encounters. Our model was designed with end-users in mind for integration within a shared decision aid. It is of paramount importance that we transparently communicate the uncertainty around estimates by explaining the cohort from which the model was derived, and include an understandable explanation of the CIs for each estimate. Of note, our model performs better than FEV1 alone in predicting survival for people with FEV1 <45% predicted. We believe that despite the wide CIs for some scenarios, we can prompt conversations about advance care planning in subsets of people with high-predicted probabilities of death. In fact, some patients may be motivated to begin discussions about advance care planning despite this large range in estimated probability of survival, even if, for example, the range is from 70% to 90%. Individual people with CF will have individual interpretations of these risk ranges when applied to themselves.<sup>26</sup> Carefully crafting communication using best-standards derived from the patient-centred communication and psychology of decision-making literature, as well as iterative testing of the interpretation of these data will be essential. Testing the acceptability of this communication, and ease of use—within a carefully designed decision aid for supporting shared decision-making—will



determine whether these models are found to be useful by patients, caregivers and their clinicians. This work is ongoing and the results of feasibility testing of a shared decision aid which incorporates these models will be separately reported.

## CONCLUSION

People with incident advanced stage CF (FEV1  $\leq$ 45%) have varying predicted 12, 24 and 36 months survival probabilities, which can be estimated using a combination of FEV1 % predicted, insurance type, the need for non-invasive ventilation and supplemental oxygen therapy, *B. cepacia* colonisation, cirrhosis, depression, renal failure requiring haemodialysis, current smoking and unclassifiable mutation class. Careful communication of these estimates and the uncertainties around them, within a well-crafted shared decision aid, has the potential to inform advance care planning and informed decision-making about treatment options.

## Author affiliations

<sup>1</sup>Division of Pulmonary Critical Care and Sleep Medicine, Department of Medicine, Northwell Health, Manhasset, New York, USA

<sup>2</sup>Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York, USA

<sup>3</sup>Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, USA

<sup>4</sup>Biostatistics, Northwell Health Feinstein Institutes for Medical Research, Manhasset, New York, USA

<sup>5</sup>Division of Nephrology, Department of Medicine, McMaster University, Hamilton, Ontario, Canada

<sup>6</sup>Division of Pulmonary, Allergy, and Critical Care, University of Pennsylvania Hospital, Philadelphia, Pennsylvania, USA

**Acknowledgements** The authors would like to thank the Cystic Fibrosis Foundation for the use of the CF Foundation Registry data to conduct this study.

**Contributors** NH, the PI of the study, generated the idea for the research study, designed the study, assisted with statistical analysis and interpretation and led the manuscript writing. MZ performed data analysis, interpretation of results and contributed to the manuscript writing. MA contributed to the data management, data analysis and interpretation of results. NK performed data analysis and results interpretation, and contributed to the manuscript writing. AM contributed to the study design and data analysis interpretation as well as manuscript writing. DH contributed to the study design and data analysis interpretation as well as manuscript writing. JW contributed to analysis and interpretation of the quantitative and qualitative data from the study, and to the writing of the manuscript. ML contributed to the data analysis and results interpretation as well as to manuscript writing.

**Funding** This study was funded by Cystic Fibrosis Foundation (HAJIZA18QI1).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data will not be readily available as it is owned by the Cystic Fibrosis Foundation. This can be accessed by the <https://www.cff.org/Research/Researcher-Resources/Patient-Registry/>

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Negin Hajizadeh <http://orcid.org/0000-0002-6404-1018>

## REFERENCES

- Ramsey BW. Management of pulmonary disease in patients with cystic fibrosis. *N Engl J Med* 1996;335:179–88.
- Corey M, Edwards L, Levison H, *et al*. Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. *J Pediatr* 1997;131:809–14.
- Rosenbluth DB, Wilson K, Ferkol T, *et al*. Lung function decline in cystic fibrosis patients and timing for lung transplantation referral. *Chest* 2004;126:412–9.
- Street RL. Aiding medical decision making: a communication perspective. *Med Decis Making* 2007;27:550–3.
- Chewning B, Bylund CL, Shah B, *et al*. Patient preferences for shared decisions: a systematic review. *Patient Educ Couns* 2012;86:9–18.
- Fried TR, Bradley EH, O'Leary J. Prognosis communication in serious illness: perceptions of older patients, caregivers, and clinicians. *J Am Geriatr Soc* 2003;51:1398–403.
- Blackhall LJ, Murphy ST, Frank G, *et al*. Ethnicity and attitudes toward patient autonomy. *JAMA* 1995;274:820–5.
- Carrese JA, Rhodes LA. Western bioethics on the Navajo reservation. benefit or harm? *JAMA* 1995;274:826–9.
- Mead EL, Doorenbos AZ, Javid SH, *et al*. Shared decision-making for cancer care among racial and ethnic minorities: a systematic review. *Am J Public Health* 2013;103:e15–29.
- Kass-Bartelmes BL, Hughes R. Advance care planning: preferences for care at the end of life. *J Pain Palliat Care Pharmacother* 2004;18:87–109.
- Charles C, Gafni A, Whelan T, *et al*. Treatment decision AIDS: conceptual issues and future directions. *Health Expect* 2005;8:114–25.
- Politi MC, Studts JL, Hayslip JW. Shared decision making in oncology practice: what do oncologists need to know? *Oncologist* 2012;17:91–100.
- McCarthy C, Dimitrov BD, Meurling IJ, *et al*. The CF-ABLE score: a novel clinical prediction rule for prognosis in patients with cystic fibrosis. *Chest* 2013;143:1358–64.
- Liou TG, Adler FR, Fitzsimmons SC, *et al*. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol* 2001;153:345–52.
- Aaron SD, Stephenson AL, Cameron DW, *et al*. A statistical model to predict one-year risk of death in patients with cystic fibrosis. *J Clin Epidemiol* 2015;68:1336–45.
- Nkam L, Lambert J, Latouche A, *et al*. A 3-year prognostic score for adults with cystic fibrosis. *J Cyst Fibros* 2017;16:702–8.
- Coriati A, Sykes J, Nkam L, *et al*. Validation of the French 3-year prognostic score using the Canadian cystic fibrosis registry. *J Cyst Fibros* 2019;18:396–8.
- MacKenzie T, Gifford AH, Sabadosa KA, *et al*. Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the cystic fibrosis Foundation patient registry. *Ann Intern Med* 2014;161:233–41.
- Fibrosis FC. Patient registry: annual data report. Available: <http://www.cff.org/UploadedFiles/research/ClinicalResearch/2011-Patient-Registry.pdf>
- Kerem E, Reisman J, Corey M, *et al*. Prediction of mortality in patients with cystic fibrosis. *N Engl J Med* 1992;326:1187–91.
- Tibshirani R. The LASSO method for variable selection in the COX model. *Stat Med* 1997;16:385–95.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
- Simon N, Friedman J, Hastie T, *et al*. Regularization paths for COX's proportional hazards model via coordinate descent. *J Stat Softw* 2011;39:1–13.
- Heagerty PJ, Lumley T, Pepe MS. Time-Dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56:337–44.



- 25 Kapnadak SG, Dimango E, Hadjiliadis D, *et al*. Cystic fibrosis Foundation consensus guidelines for the care of individuals with advanced cystic fibrosis lung disease. *J Cyst Fibros* 2020;19:344–54.
- 26 Hajizadeh N, Basile MJ, Kozikowski A, *et al*. Other ways of knowing. *Med Decis Making* 2017;37:216–29.

## Appendix 1

### Estimation of 12, 24, and 36 month Survival Probabilities:

We first estimated the baseline survivor function from the Cox model and applied the derived hazard function corresponding to  $X$  (the covariate vector). The following notation is used:  $Q(t;s)$  denotes the baseline probability of surviving 't' months given that the patient has survived through time 's'. For a patient with covariate vector  $X$ ,  $S(b;a,X)$  denotes the probability of surviving an additional 'b' months, given that the patient survived to time 'a'. The baseline survivor function, 'Q', was internally computed in R by using the `fit.glmnet` function in `glmnet` with all covariates set to zero (method of Paul Allison<sup>1</sup>). Using this approach,  $S(b;a,X)$  can then be computed for any configuration of  $X$  by simply computing  $S(b;a,X) = [Q(b;a)]^{\exp(BX)}$

Estimation of 12, 24, and 36 month Survival Probabilities for those with incident advanced stage CF (T12): The estimated baseline 12, 24 and 36-month survival probabilities, for those from the time of incident advanced stage CF diagnosis is  $Q(12;0)=0.780$ ,  $Q(24;0)=0.573$ , and  $Q(36;0)=0.412$  respectively. For example,  $S(12;0,X)$ , the probability of 12-month survival from incident severe CF, for a patient with covariates  $X=(X_1,X_2,\dots,X_7)$  and estimated beta coefficients  $B=(B_1,B_2,\dots,B_7)$ , is estimated by  $0.880^{\exp(BX)}$  (we use the following notation: for a patient with covariate vector  $X$ ,  $S(b;a,X)$  denotes the conditional probability of surviving an *additional* b months, given that the patient survived to time 'a'). For example, as outlined in Table 3a, consider a patient with the following "X" factors (covariates): FEV1 (%) predicted: 40, Private Insurance only, No smoking, No non-invasive ventilation, No oxygen Therapy, No *Burkholderia* species, No liver disease, cirrhosis, No depression, No Renal failure requiring dialysis, Mutation class: 1-3, Number of Pulmonary Exacerbations in the year preceding advanced stage CF diagnosis: 0, Baseline Lung Transplant Evaluation status: not pertinent. The estimated probability of 12-month survival is  $S(12;0,X) = 0.780^{\exp(-2.20)} = 0.973$ , or, 88.3%. We report the probabilities of survival to 12 months, 24 months and 36 months from the time of incident advanced stage CF diagnosis (i.e.  $S(12;0,X)$ ,  $S(24;0,X)$ ,  $S(36;0,X)$ , respectively). Examples of combinations of covariables are shown in Table 4a.

Estimation of 12, 24, and 36 month Survival Probabilities for those surviving to 12 months after incident advanced stage CF (T12): The estimated baseline 12, 24 and 36-month survival probabilities, for those who have survived 12 months from the time of incident advanced stage CF diagnosis is  $Q(12;12)=0.924$ ,  $Q(24;12)=0.861$ , and  $Q(36;12)=0.796$  respectively. Thus,  $S(12;12,X)$ , the conditional probability of 12-month survival for a patient who survived the first 12 months from incident severe CF, with covariates  $X=(X_1,X_2,\dots,X_7)$  and estimated beta coefficients  $B=(B_1,B_2,\dots,B_7)$ , is estimated by  $0.924^{\exp(BX)}$ . Examples of combinations of covariables are shown in Table 4b. The 95% confidence of the predicted survival probabilities were obtained using non-parametric bootstrap (27).