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[1]). 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 $X=(X_1,X_2,\ldots,X_7)$ and estimated beta coefficients $B=(B_1,B_2,\ldots,B_7)$, 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 Q36;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 $X=(X_1,X_2,\ldots,X_7)$ and estimated beta coefficients $B=(B_1,B_2,\ldots,B_7)$, 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).