

# Explaining the Sex Effect on Survival in Cystic Fibrosis: a Joint Modeling Study of UK Registry Data

David Taylor-Robinson,<sup>a</sup> Daniela K. Schlüter,<sup>a</sup> Peter J. Diggle,<sup>b</sup> and Jessica K. Barrett<sup>c,d</sup>

**Background:** Male sex is associated with better lung function and survival in people with cystic fibrosis but it is unclear whether the survival benefit is solely due to the sex-effect on lung function.

**Methods:** This study analyzes data between 1996 and 2015 from the longitudinal registry study of the UK Cystic Fibrosis Registry. We jointly analyze repeated measurements and time-to-event outcomes to assess how much of the sex effect on lung function also explains survival. These novel methods allow examination of association between percent of forced expiratory volume in 1 second (%FEV1) and covariates such as sex and genotype, and survival, in the same modeling framework. We estimate the probability of surviving one more year with a probit model.

**Results:** The dataset includes 81,129 lung function measurements of %FEV1 on 9,741 patients seen between 1996 and 2015 and captures 1,543 deaths. Males compared with females experienced a more gradual decline in %FEV1 (difference 0.11 per year 95% confidence interval [CI] = 0.08, 0.14). After adjusting for confounders, both overall level of %FEV1 and %FEV1 rate of change are associated

with the concurrent hazard for death. There was evidence of a male survival advantage (probit coefficient 0.15; 95% CI = 0.10, 0.19) which changed little after adjustment for %FEV1 using conventional approaches but was attenuated by 37% on adjustment for %FEV1 level and slope in the joint model (0.09; 95% CI = 0.06, 0.12).

**Conclusions:** We estimate that about 37% of the association of sex on survival in cystic fibrosis is mediated through lung function.

**Keywords:** Cystic fibrosis; Joint-modeling; Registry

(*Epidemiology* 2020;31: 872–879)

Male sex has been identified as a positive prognostic factor in cystic fibrosis.<sup>1–3</sup> An apparent effect of sex on morbidity and mortality in cystic fibrosis, with males having better outcomes, has been a common finding in large epidemiologic studies, first suggested in US centers,<sup>4</sup> and then confirmed in US population-level registry studies.<sup>5,6</sup> There have been similar findings in the UK,<sup>7</sup> with Barr et al in the UK<sup>8</sup> suggesting that despite overall improved survival in the 21st century, females continue to be more likely to die below the median age of death compared with males, a pattern that has persisted since the 1960s. There has been recent debate about the sex gap, suggesting that this may be narrowing over time as a result of improving treatment.<sup>9,10</sup> In terms of use of health services in cystic fibrosis, a large study in Canada further demonstrated decreased risk of hospitalization in males.<sup>11</sup> Our previous study of the UK Cystic Fibrosis Registry shows that the sex difference in cystic fibrosis outcomes is clearly apparent in the UK cystic fibrosis population<sup>12</sup> and this is also shown in a recent survival analysis using UK data.<sup>13,14</sup> The cause of the sex gap in survival remains unclear. Lung function, as measured by the percent predicted forced expiratory volume in 1 second (%FEV1) is commonly used as a measure of disease severity in cystic fibrosis, and has been shown to be related to survival. One explanation for the sex gap in cystic fibrosis survival is that this is explained by worse lung function in females. For instance, some studies suggest that females may be more likely to become colonized with *Pseudomonas aeruginosa* leading to lung damage at an earlier age,<sup>15</sup> and this may be related to the effect of estrogen.<sup>16</sup> A recent review of the sex gap in cystic fibrosis has suggested that the finding of lower female survival is evident in most studies, and that evidence to suggest closure

Submitted July 10, 2019; accepted July 27, 2020.

From the <sup>a</sup>Department of Public Health and Policy, University of Liverpool, Liverpool, United Kingdom; <sup>b</sup>CHICAS, Lancaster Medical School, Lancaster University, Lancaster, United Kingdom; <sup>c</sup>Department of Public Health and Primary Care, University of Cambridge, United Kingdom; and <sup>d</sup>MRC Biostatistics Unit, University of Cambridge, United Kingdom.

D.T.R., D.K.S., and P.J.D. are supported by a UK cystic fibrosis Trust Strategic Research Centre grant. J.K.B. is supported by an MRC Unit programme. David Taylor-Robinson is funded by the MRC on a Clinician Scientist Fellowship (MR/P008577/1).

D.T.R., D.K.S., and P.J.D. are supported by a UK cystic fibrosis Trust Strategic Research Centre grant. The other author has no conflicts to report.

**SDC** Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article ([www.epidem.com](http://www.epidem.com)).

D.T.R., P.J.D., and J.K.B. are guarantors. D.T.R., P.J.D., and J.K.B. conceived and designed. D.T.R., D.K.S., P.J.D., and J.K.B. involved in data collection. D.T.R., D.K.S., P.J.D., and J.K.B. responsible for analysis and interpretation. D.T.R., P.J.D., and J.K.B. contributed to drafting the manuscript for important intellectual content.

Code for replication: The code for fitting this model is available via the link <https://github.com/Jessbarrett/CysticFibrosisJM>.

Correspondence: David Taylor-Robinson, Department of Public Health and Policy, Farr Institute@HeRC, University of Liverpool, Liverpool, United Kingdom. E-mail: [david.taylor-robinson@liverpool.ac.uk](mailto:david.taylor-robinson@liverpool.ac.uk).

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ISSN: 1044-3983/20/3106-0872

DOI: 10.1097/EDE.0000000000001248

of the gap in recent cohorts is less convincing than the data supporting its continued existence. The review does suggest that in cohorts of adults with late diagnosis, and conditional on survival to age 40 years, the sex gap appears to narrow or even be reversed.<sup>3</sup>

In this study, we aim to quantify how aspects of an individual cystic fibrosis patient's longitudinal profile of lung function are related to their survival prognosis; and to decompose the impact of sex on these joint outcomes. Joint modeling approaches are potentially of great utility in the context of studying outcomes in cystic fibrosis patients.<sup>17</sup> Survival is of central interest and analyses often seek to adjust for lung function as a time-varying covariate, which we know is measured imprecisely with clinically significant measurement error. Furthermore, the dynamics of lung function decline are also of interest, but there is potentially informative drop-out due to the direct link between lung function and survival prognosis. Together, these properties of cystic fibrosis data (measurement error and dropout) mean that separate analysis of repeated measurement and survival outcomes is potentially inefficient, because it does not exploit the dependence between the repeated measurement process and the hazard for survival, and leads to biased estimation of the association between the two because it ignores measurement error.<sup>18</sup> Joint modeling of lung function and survival offers an approach to address all of these issues.

Joint modeling has been applied to cystic fibrosis data in a few previous studies. The first of these was a center-level study by Schluchter and colleagues that modeled longitudinal FEV1 and survival simultaneously for a cohort of  $\Delta F508$  homozygous patients, but this study did not explore the sex effect.<sup>17</sup> Subsequently, cystic fibrosis data have been used to develop methods for joint modeling, including an approach that we previously developed.<sup>19</sup> In this study, we apply this novel approach for the joint modeling of lung function and survival and contrast this to commonly used approaches to adjusting for time-varying covariates in survival analyses. We use the joint model to test the hypothesis that the survival advantage for males is explained by the effect of sex on lung function.

## METHODS

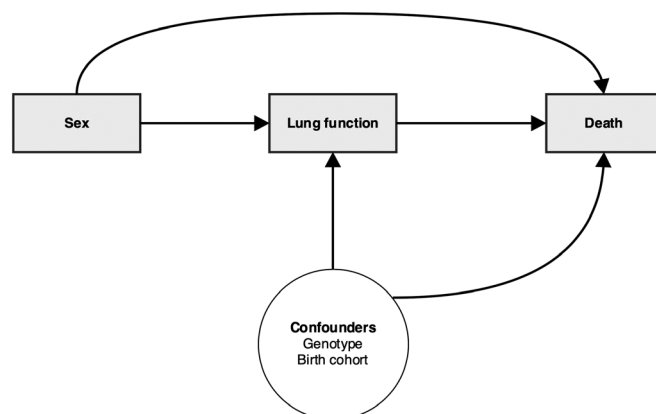
We undertook a longitudinal retrospective cohort study of individuals in the UK Cystic Fibrosis Registry, which records longitudinal health data on all people with cystic fibrosis in England, Wales, Scotland, and Northern Ireland. All UK cystic fibrosis centers and clinics routinely collect data in a standardized fashion. When patients with cystic fibrosis attend a new center in the UK, they or their parents consent to collection and storage of information on the patient's health and treatment in the Cystic Fibrosis Registry.<sup>20</sup> In the UK, cystic fibrosis patients are seen in the outpatient clinic for a comprehensive annual review, including evaluation of clinical status and pulmonary function. The UK Cystic Fibrosis Registry is supported and coordinated by the UK Cystic fibrosis

Trust. In the UK, the Registry is estimated to capture almost all of the cystic fibrosis population; any consenting patients attending National Health Service clinics will have annual data routinely collected into the database, 89% of whom have a "complete" dataset capturing key clinical parameters over time. The registry is carefully managed and curated to a high level of data quality, and is therefore ideally suited to the study of cystic fibrosis outcomes.<sup>20</sup>

## Primary Outcome and Covariates

Our directed acyclic graph which informed the analysis, identifying variables in the minimally sufficient set of adjustments, is shown in Figure 1. The primary outcomes were %FEV1, as per other studies that have explored the cystic fibrosis sex gap in outcomes,<sup>5</sup> and survival. Pulmonary function tests were measured annually at the review visit, and performed according to international recommendations, measuring forced expiratory volume in 1 second, expressed as a percentage of predicted values for sex, age, height, and ethnicity using the global lung function initiative reference equations.<sup>21</sup> We restricted the analysis to white patients under the age of 40 at last follow-up, with at least one lung function measurement between the start of 1996 and the end of 2015. We chose to apply an upper age limit to the analysis since the female sex gap has been shown to be present up to this point; and we have previously shown that random intercept and slope models make unrealistic assumptions when applied over long periods.<sup>22</sup> About 97% of people with cystic fibrosis in the UK are white. Non-white patients tend to have worse outcomes, but the numbers are small in the UK, restricting power to demonstrate subgroup effects.<sup>12</sup>

The primary exposure of interest was male sex. The time metameter was patient age at clinic visit at which the %FEV1 measure was taken. Other covariates in the analysis were genotype coded as the number of delta F508 alleles (0, 1, or 2) and dichotomized into 2 F508 alleles versus 0 or 1 alleles or not typed, and birth year which was treated as a



**FIGURE 1.** Directed acyclic graph for the effect of sex on key outcomes in cystic fibrosis. We aim to use a joint model to test the hypothesis that there is a direct effect of sex on survival.

continuous covariate and centered at the mean value (approximately 1986) to capture cohort effects.

## Statistical Analysis

Full details of the joint modeling approach are provided in the supplementary eAppendix 1; <http://links.lww.com/EDE/B714>. We applied the method developed by Barrett et al which allows exact likelihood inference for a wide range of random-effect specifications, and the code for fitting this model is available via the link (<https://github.com/Jessbarrett/CysticFibrosisJM>). Repeated %FEV1 measures on individuals are correlated, and this must be accommodated to obtain valid inferences. Furthermore, lung function is related to survival. Thus, repeated measurements of %FEV1 and survival were modeled jointly using shared random effects to account for the interdependence of the two processes.<sup>19</sup> The submodel for %FEV1 adjusted for the patient's age at measurement, birth year, sex, and number of F508 alleles. Exploratory plots of the data are shown in eAppendix2; <http://links.lww.com/EDE/B714>. Informed by these, we approximated time-trends with a quadratic time function, to accommodate nonlinear change over time.<sup>22</sup> Interactions were included between the linear time variables in age and birth year and all other covariates. The birth year was included to account for cohort effects and a survivor effect arising from left truncation of the data at the start of follow-up. The intercept (level of %FEV1 at age five) and age effect (annual linear change in %FEV1) were treated as normally distributed, correlated random effects to allow for individual intercepts and slopes. The random effects in the longitudinal model capture other sources of unmeasured heterogeneity not captured by the fixed effects.

Subjects entered the cohort at different ages so that patients contributed person-years to the analysis only at ages corresponding to their actual ages during the study period.

The probability of surviving one year was modeled using a probit link function. The survival submodel was adjusted for age halfway through the year, birth year, sex, and F508 alleles (dichotomized as previously). In addition, the survival submodel depended on a patient's %FEV1 value halfway through the year and their %FEV1 rate of change, as estimated by the longitudinal %FEV1 model including the random effects. The time-to-event submodel is specified on a discrete-time scale modeling the conditional probability of surviving one year given that you have survived to the start of the year, whereby a positive coefficient means that an increase in the predictor leads to an increase in the predicted probability of survival. The effect of %FEV1 level and %FEV1 rate of change on survival were also expressed as hazard ratios (see the statistical eAppendix1; <http://links.lww.com/EDE/B714> for details of estimating hazard ratios from our survival model). Note that for the hazard ratios, the direction of effect is reversed, i.e. a hazard ratio >1 means that an increase in the predictor gives a decreased probability of survival.

We assess whether there is a direct association between sex and survival according to the directed acyclic graph as shown in Figure 1, after accounting for lung function. Our coefficient of interest in the joint model is the sex effect in the survival submodel, after adjustment for %FEV1 rate of change and overall level of %FEV1 (for full algebraic details see eAppendix 1; <http://links.lww.com/EDE/B714>). To demonstrate the utility of using a joint modeling approach to test our hypothesis, we assessed how the association between sex and survival changed depending on the modeling approach. Once we fitted the joint model to the data, we compared the association between sex and survival with that in a standard probit survival model, first without any adjustment for %FEV1, and then with %FEV1 added as a baseline time-invariant covariate (i.e. %FEV1 at first visit), and finally as a time-varying covariate. We estimated all model parameters by maximum likelihood and used generalized likelihood ratio statistics to compare nested models, and Wald statistics to test hypotheses about model parameters. We plotted residual diagnostics for the longitudinal and survival submodels, and an empirical variogram to check our model fit.

## Ethics

National Health Service research ethics approval (Huntingdon Research Ethics Committee 07/Q0104/2) has been granted for the collection of data into the UK database. The cystic fibrosis Trust database committee approved the use of anonymized data in this study.

## RESULTS

### Population Characteristics

The dataset contained 81,129 lung function measures on 9,741 patients between 1996 and 2015 in the UK and captured 1,543 deaths. The median number of %FEV1 measures per person was 8 (range 1–25). About 87% of individuals had three or more follow-up measures with a total of 96,598 person-years of follow-up. The baseline characteristics of the population, stratified by sex, are shown in Table 1.

### Associations of Covariates with Lung Function Trajectory

We explored the effect of covariates on %FEV1. Results from the %FEV1 submodel are shown in Table 2. Random effect parameter estimates are reported in eAppendix2; <http://links.lww.com/EDE/B714>. There was a borderline difference between males and females in the level of %FEV1 at age 5 with the difference in intercept of 0.90 (95% CI = 0.01, 1.80). %FEV<sub>1</sub> initially declined at a rate of -1.52 (95% CI = -1.58, -1.45) percentage points per year for females. Males experienced a more gradual decline in %FEV1 (difference 0.11 per year 95% CI = 0.08, 0.14). Lung function declined at a faster rate for patients with 2 F508 alleles compared with those with 0 or 1 F508 alleles (-0.35 percentage points per year 95% CI = -0.45, -0.25).

**TABLE 1.** Characteristics of Study Population by Sex: UK Cystic Fibrosis Registry

	Female	Male	All
N (%)	4605 (47)	5136 (53)	9741
Observation, n (%)	37849 (47)	43280 (53)	81129
Observations per patient, mean (SD)	8.2 (4.7)	8.4 (4.8)	8.3 (4.8)
Deaths, n (%)	813 (53)	730 (47)	1543
Genotype			
No. delta 508: 2, n (%)	2361 (51)	2752 (54)	5113 (53)
No. delta 508: 1, n (%)	1718 (37)	1782 (35)	3500 (36)
No. delta 508: 0, n (%)	327 (7.1)	354 (6.9)	681 (7.0)
Missing	199 (4.3)	248 (4.8)	447 (4.6)
Birth cohort, n (%)			
<1960	3 (0.1)	4 (0.1)	7 (0.1)
1960–1964	68 (1.5)	115 (2.2)	183 (1.9)
1965–1969	141 (3.1)	199 (3.9)	340 (3.5)
1970–1974	257 (5.6)	335 (6.5)	592 (6.1)
1975–1979	365 (7.9)	473 (9.2)	838 (8.6)
1980–1984	587 (13)	686 (13)	1273 (13)
1985–1989	668 (15)	772 (15)	1440 (15)
1990–1994	722 (16)	749 (15)	1471 (15)
>1995	1794 (39)	1803 (35)	3597 (37)
Age at entry, yrs, mean (SD)	19.2 (9.0)	20.0 (9.4)	19.6 (9.2)
Age at diagnosis, yrs, mean (SD) (missing, n = 118)	3.0 (6.5)	3.0 (6.8)	3.0 (6.7)

### Associations of Covariates with Survival

Table 3 shows the results from the survival models. The joint model (column 1) shows that higher overall levels of %FEV1 and a more gradual %FEV1 decline were associated with improved survival. Figure 2 provides a visual illustration of the relationship between lung function and survival for individuals selected from the population with different %FEV1 trajectories (different random effects). For example, for 20-year old males born in 1980 with two F508 alleles and population-average %FEV1 level and %FEV1 decline, a 10-unit lower level of %FEV1 is associated with an increase in the concurrent hazard of death (HR 2.26 CI = 2.13, 2.41). Furthermore, those with a one-unit-per-year decline compared with no decline in %FEV1 have over a three-fold higher hazard of death for every year that passes (HR 3.67 CI = 3.31, 4.07).

Figure 3 visualizes the association of sex with lung function and survival in the joint model. Males have a more favorable %FEV1 and survival trajectory. Note that the estimated sex effects in Figure 3 are population-averaged effects; that is, they describe average values of %FEV1 for subpopulations of individuals sharing the same explanatory characteristics, rather than for any one individual.

**TABLE 2.** Joint Model Results for the %FEV1 Submodel. Fixed Effects Estimates of Association of Covariates on Forced Expiratory Volume in 1 second as a Percentage of Predicted (%FEV1)

	Estimate (95% CI)
Intercept at age 5 years	88 (87, 89)
Age <sup>a</sup>	-1.5 (-1.6, -1.4)
Age squared	0.013 (0.011, 0.015)
Birth year	0.22 (0.15, 0.29)
Male	0.90 (0.01, 1.8)
F508 alleles: 2 vs 0, 1 or not typed	-0.080 (-1.17, 1.01)
Age × male	0.11 (0.08, 0.14)
Age × F508 alleles: 2 vs 0, 1 or not typed	-0.35 (-0.45, -0.25)
Birth year × male	0.059 (-0.027, 0.15)
Birth year × F508 alleles: 2 vs 0, 1 or not typed	-0.0040 (-0.067, 0.059)

<sup>a</sup>Age term corresponds to %FEV1 slope. The age × male term represents the age by sex interaction, i.e. the difference in slope for males compared with females  
%FEV1 indicates percent forced expiratory volume.

### Association Between Female Sex and Survival Adjusted for Lung Function

In Table 3, we compare the results of four possible survival models that adjust for %FEV1 in various ways. Contrasting the association of sex with survival across these models, in all three of the standard survival models there is a strong association between sex and survival in models without adjustment for %FEV1, and with adjustment for %FEV1 either as a baseline or as a time-varying covariate (effect sizes 0.15, 0.13, 0.14, i.e. males have better survival than females). By contrast, we observe less association between sex and survival after adjusting for %FEV1 (level and rate of decline) in the joint model, with an effect size attenuated by 37% to 0.09 (95% CI = 0.06, 0.12). Table 3 also shows that the genotype effect is reversed once we adjust for longitudinal lung function in the joint model. Residual diagnostics did not raise any concerns about model fit (eAppendix2; <http://links.lww.com/EDE/B714>).

### DISCUSSION

We apply a novel joint modeling approach, for the first time, to show that about 40% of the association of female sex on survival in cystic fibrosis is explained by the effect of sex on lung function. Both increased rates of decline in lung function, and a decreased overall level of lung function, are associated with an increased risk of death. The strength of this analysis is the use of a joint modeling approach that leads to more robust estimation of both survival estimates and the rate of lung function decline. Our joint modeling approach allows exploration of how aspects of an individual cystic fibrosis patient's longitudinal profile of %FEV1 are related to their survival prognosis. A key advantage of the joint model is that it estimates the relationship between characteristics of the "true" error-free underlying %FEV1 trace and survival. Our

**TABLE 3.** Parameter Estimates (95% CI) From the Probit Survival Models. A Positive Coefficient Means that an Increase in the Predictor Leads to an Increase in the Predicted Probability of Survival

	Joint model	No %FEV1	Baseline %FEV1	Time-varying %FEV1
Intercept	1.6 (1.5, 1.7)	2.4 (2.3, 2.4)	1.3 (1.1, 1.4)	0.37 (0.23, 0.50)
%FEV1 (per 10 units)	0.28 (0.26, 0.29)		0.18 (0.17, 0.19)	0.32 (0.31, 0.33)
%FEV1 slope	0.45 (0.43, 0.48)			
Age-5 (per 10 yrs)	-0.042 (-0.059, -0.026)	-0.046 (-0.089, -0.002)	-0.12 (-0.17, -0.08)	0.059 (0.007, 0.111)
Birth year (per 10 yrs)	0.26 (0.22, 0.30)	0.16 (0.12, 0.21)	-0.073 (-0.120, -0.025)	0.039 (-0.011, 0.090)
Male	0.092 (0.062, 0.122)	0.15 (0.10, 0.19)	0.13 (0.09, 0.18)	0.14 (0.09, 0.19)
F508 alleles: 2 vs 0, 1 or not typed	0.14 (0.03, 0.25)	-0.13 (-0.17, -0.09)	-0.10 (-0.14, -0.06)	-0.042 (-0.089, 0.006)

%FEV1 indicates percent forced expiratory volume.

analysis here suggests that measurement error – in conventional approaches to adjusting for %FEV1 in survival analysis – has led to previous estimates of a substantial association between sex and survival. This overestimate was attenuated once error-free %FEV1 was taken into account in the joint model.

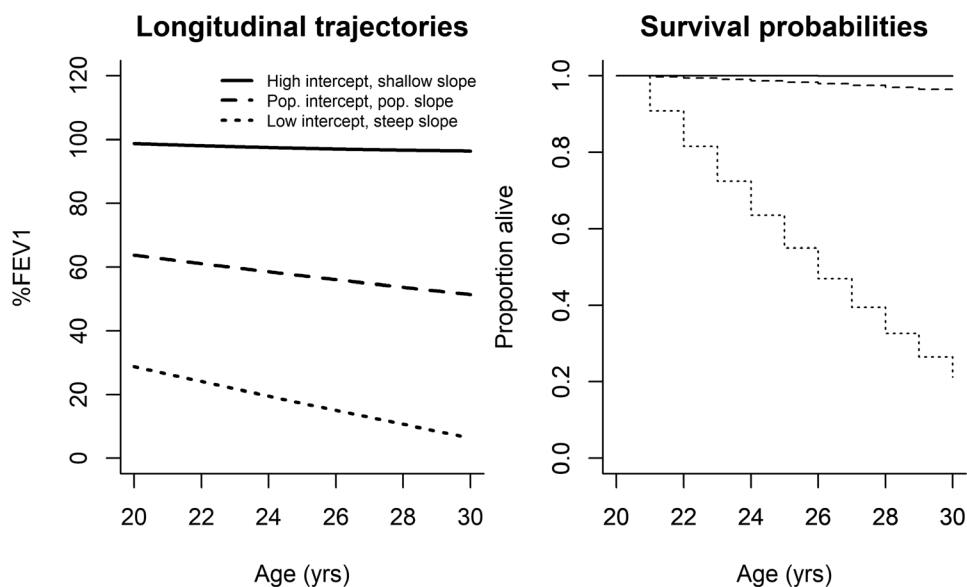
Many studies have explored survival in cystic fibrosis,<sup>1,2,13</sup> and notable among these are the large studies from North America that have used Cox regression to estimate the effect of various covariates on survival chances in registry populations.<sup>5,6,23</sup> A large number of factors have been identified that may influence survival, with female sex identified most commonly as a risk factor. Other influences include poor respiratory function and risk factors for poor lung function such as *P. aeruginosa* infection status, homozygous delta F508 status, or heterozygous nondelta F508 status, non-white ethnicity, and low income. Many studies have investigated the impact of baseline time-invariant factors on survival, including a recent study using UK data by Keogh and colleagues which showed a clear sex gap, with worse survival for female patients.<sup>13</sup> Where studies have investigated time varying predictors such as lung function, these have been included in survival models as current values. To our knowledge, our study is one of the first to quantify the association of rate of lung function decline with survival using a joint modeling approach. Another recent study by Keogh et al used landmarking to predict survival in the UK population, and in a two-stage approach used fitted values from a longitudinal mixed-effects model for %FEV1 to estimate current values of %FEV1 for inclusion in the second-stage landmarking analysis.<sup>14</sup> The study found that current %FEV1 was the strongest predictor of survival. As in our joint model, this two-stage landmarking approach could also be used to estimate the association of the recent rate of decline of %FEV1 with survival.

There is a large literature on lung function decline in people with cystic fibrosis. Konstan et al have undertaken the largest studies to date of %FEV1 in both pediatric and adult cohorts,<sup>24–26</sup> using a mixed-effects regression approaches, to show that higher baseline %FEV1, *P. aeruginosa* colonization, female sex, and poor nutritional status were amongst the factors associated with a greater decline in lung function in

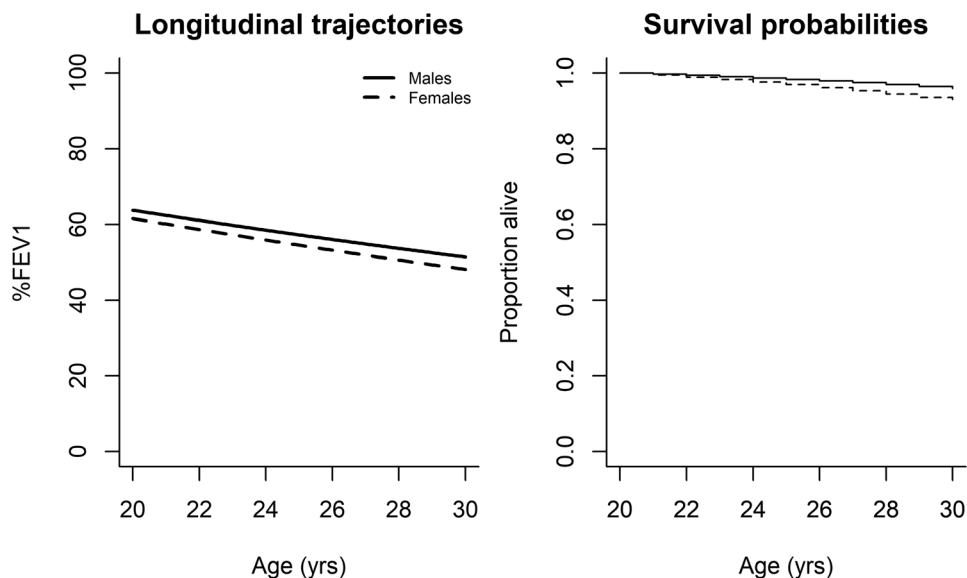
people with cystic fibrosis.<sup>24</sup> Our analysis contributes to this literature by explicitly quantifying how longitudinal changes in lung function are related to survival. Both an increased rate of decline in lung function, and a decreased overall level of lung function are associated with an increased risk of death in people with cystic fibrosis.

Large national cystic fibrosis patient registry data in different countries have shown that survival for cystic fibrosis females was less than that for matched cystic fibrosis males, in an apparent sex gap. The etiology of this gap is poorly understood, with increased mortality in females seen even after correcting for lung function, suggesting the explanation is likely to be multifactorial. Our analysis adds to this literature, estimating that about 40% of the effect of sex on survival is explained by the impact of sex on lung function decline, which in turn influences survival. There are known effects of sex on lung function and acquisition of *P. aeruginosa* that we have previously identified in this dataset.<sup>12</sup> Others have suggested alternative biologic reasons, including the impact of estrogen and increased occurrence of cystic fibrosis-related diabetes (CFRD).<sup>3,8</sup> Social explanations proposed may relate to gender roles, such as a possible propensity to less exercise in childhood in girls, and an increased tolerance of poor nutritional status in adolescent girls with cystic fibrosis, fuelled by the societal pressure to appear thin.<sup>27</sup>

Our analysis suggests that the sex gap may be partly explained by worse lung function in females. We suggest that standard approaches for adjusting for lung function in survival analysis may lead to insufficient adjustment, which may explain the difference between our results, and other large epidemiologic studies of cystic fibrosis. Because lung function is measured with significant error, the associations with other covariates in the analysis may be confounded by residual effects of lung function. In our joint model, which more precisely adjusts for underlying lung function, differences in lung function explain about 40% of the sex difference in mortality. For instance, one of the largest studies to explore the gender gap in cystic fibrosis was Rosenfeld and colleagues' analysis of US registry data,<sup>5</sup> which showed a significant association of sex on survival and lung function. The authors show that



**FIGURE 2.** Estimated longitudinal trajectories and survival curves for 20-year-old males with varying intercepts and slopes, born in 1980, with 2 F508 alleles.



**FIGURE 3.** Estimated longitudinal trajectories and survival curves for males and females age 20, born in 1980 with 2 F508 alleles. Effect of sex on forced expiratory volume in 1 second as a percentage of predicted (%FEV1) and survival in the final joint model.

pulmonary function was the only risk factor that explained a portion of the observed gender-related difference in survival. Among the subjects 1-20 years of age, females had a hazard ratio of 1.7 in the unadjusted Cox regression analysis, but this was attenuated to 1.5 while adjusting for %FEV1 as a time-varying covariate. The authors suggest that differences in pulmonary function did appear to explain a small portion of the excess female mortality, but no other factor further accounted for the gender gap. A more recent large study of US registry data identified a survival advantage for males compared with females, with a 19% (CI = 13%, 24%) lower adjusted risk for death in males as compared with females,<sup>28</sup> and the authors further highlight that the reasons for this are not well-understood. However, this analysis did not adjust for lung function

at a time-varying covariate but focused on adjustment for baseline factors at the time of diagnosis, which can be used by clinicians at diagnosis to inform discussions about patient prognosis.

To our knowledge, this is the first study to explore the sex effect in cystic fibrosis using joint modeling approaches, though previous studies have used cystic fibrosis data to develop joint modeling methods.<sup>19</sup> Key strengths of this study include the population-wide coverage of the UK Cystic Fibrosis Registry and the high quality of the data.<sup>20</sup> Although we have not presented our analysis as formal mediation analysis, the steps we have undertaken map onto the Baron and Kenny<sup>29</sup> steps for mediation analysis, subject to a number of assumptions (see eAppendix 2; <http://links.lww.com/EDE/B714>).

Newer methods for mediation analysis based on the counterfactual framework, and software to implement them, are developing rapidly, and further work is warranted to understand how joint models can be used within the potential outcomes framework. Such approaches could be usefully applied to range of mediation questions in cystic fibrosis epidemiology. For example, while not the main focus of our analysis here, the reversal of the genotype effect in our joint model suggests that heterozygotes for del508, compared with those with one or no del508 alleles, have better survival after adjustment for lung function. Further analyses could explore decomposing the sex effect on survival through a potential effect of sex on *P. aeruginosa* acquisition which is known to impact lung function decline.

There are a number of limitations to our analysis: First, it relies on retrospective, routinely collected data. Second, our joint modeling approach assumes that %FEV1 measurements are independent of survival given an individual's random-effect values, which may not be appropriate. Third, we did not explore the full range of mediating pathways through which sex may impact survival. Our main hypothesis related to the sex term in the survival submodel, and we did not seek to adjust for other potential downstream mediators of the association between sex and survival in the survival submodel, such as CFRD and infection status indicators, instead adjusting for slope and overall level of lung function as captured in our longitudinal submodel. Whilst the random effects included in the longitudinal component of our joint model capture residual underlying baseline heterogeneity they are unlikely to adequately capture time-varying effects such as infections. The impact of infection on survival, however, is likely to be largely mediated through impacts on lung function. A recent study has illustrated how left truncation can lead to biased estimates in the context of joint models.<sup>30</sup> We aimed to limit the impact of left truncation by including the birth year as a covariate in the longitudinal and survival submodels. The effect of birth year here is driven by both cohort effects and left truncation. Another limitation of our study is that it is debatable how transplanted individuals should be handled in a joint model.<sup>14</sup> For the purposes of this analysis, we included post-transplant FEV1 measures and deaths, but further work is needed to understand how best to take account of transplantation and post-transplant survival in the context of a joint modeling analysis.

In summary, we have applied the Barrett et al<sup>19</sup> approach to joint modeling in cystic fibrosis to address the question of how the sex effect on survival is explained by lung function. Our analysis suggests that if the lung function gap between males and females can be narrowed, this should also narrow the survival gap. Our analysis approach can be applied to similar etiologic questions in longitudinal Cystic Fibrosis Registry data and can be used to more accurately adjust for time-varying covariates measured with error, such

as lung function, in survival analyses in cystic fibrosis and other conditions.

## ACKNOWLEDGMENTS

We thank the UK Cystic Fibrosis Trust for access to the UK Cystic Fibrosis Registry and all the Centre Directors for the input of data to the registry. Elaine Gunn assisted with access to the data. JKB was supported by the MRC (grant number G0902100 and unit programme number MC\_UU\_00002/5). DTR is funded by the MRC on a Clinician Scientist Fellowship (MR/P008577/1). The work was supported by the Strategic Research Centre "EpiNet: Harnessing data to improve lives" funded by the Cystic Fibrosis Trust.

## REFERENCES

- Salvatore D, Buzzetti R, Baldo E, et al. An overview of international literature from cystic fibrosis registries. Part 4: update 2011. *J Cyst Fibros*. 2012;11:480–493.
- Buzzetti R, Salvatore D, Baldo E, et al. An overview of international literature from cystic fibrosis registries: 1. Mortality and survival studies in cystic fibrosis. *J Cyst Fibros*. 2009;8:229–237.
- McIntyre K. Gender and survival in cystic fibrosis. *Curr Opin Pulm Med*. 2013;19:692–697.
- Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. *N Engl J Med*. 1992;326:1187–1191.
- Rosenfeld M, Davis R, FitzSimmons S, Pepe M, Ramsey B. Gender gap in cystic fibrosis mortality. *Am J Epidemiol*. 1997;145:794–803.
- O'Connor GT, Quinton HB, Kahn R, et al.; Northern New England Cystic Fibrosis Consortium. Case-mix adjustment for evaluation of mortality in cystic fibrosis. *Pediatr Pulmonol*. 2002;33:99–105.
- Dodge JA, Lewis PA, Stanton M, Wilsher J. Cystic fibrosis mortality and survival in the UK: 1947–2003. *Eur Respir J*. 2007;29:522–526.
- Barr HL, Britton J, Smyth AR, Fogarty AW. Association between socioeconomic status, sex, and age at death from cystic fibrosis in England and Wales (1959 to 2008): cross sectional study. *BMJ*. 2011;343:d4662.
- Viviani L, Bossi A, Assael BM; Italian Registry for Cystic Fibrosis Collaborative Group. Absence of a gender gap in survival. An analysis of the Italian registry for cystic fibrosis in the paediatric age. *J Cyst Fibros*. 2011;10:313–317.
- Verma N, Bush A, Buchdahl R. Is there still a gender gap in cystic fibrosis? *Chest*. 2005;128:2824–2834.
- Stephenson A, Hux J, Tullis E, Austin PC, Corey M, Ray J. Higher risk of hospitalization among females with cystic fibrosis. *J Cyst Fibros*. 2011;10:93–99.
- Taylor-Robinson DC, Smyth RL, Diggle PJ, Whitehead M. The effect of social deprivation on clinical outcomes and the use of treatments in the UK cystic fibrosis population: a longitudinal study. *Lancet Respir Med*. 2013;1:121–128.
- Keogh RH, 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*. 2018;17:218–227.
- Keogh RH, Seaman SR, Barrett JK, Taylor-Robinson D, Szczesniak R. Dynamic prediction of survival in cystic fibrosis: a landmarking analysis using UK patient registry data. *Epidemiology*. 2019;30:29–37.
- Rosenfeld M, Emerson J, McNamara S, et al.; EPIC Study Group. Risk factors for age at initial *Pseudomonas* acquisition in the cystic fibrosis epic observational cohort. *J Cyst Fibros*. 2012;11:446–453.
- Chotirmall SH, Smith SG, Gunaratnam C, et al. Effect of estrogen on *Pseudomonas* mucoidy and exacerbations in cystic fibrosis. *N Engl J Med*. 2012;366:1978–1986.
- Schluchter MD, Konstan MW, Davis PB. Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. *Stat Med*. 2002;21:1271–1287.

18. Asar Ö, Ritchie J, Kalra PA, Diggle PJ. Joint modelling of repeated measurement and time-to-event data: an introductory tutorial. *Int J Epidemiol.* 2015;44:334–344.
19. Barrett J, Diggle P, Henderson R, Taylor-Robinson D. Joint modelling of repeated measurements and time-to-event outcomes: flexible model specification and exact likelihood inference. *J R Stat Soc Series B Stat Methodol.* 2015;77:131–148.
20. Taylor-Robinson D, Archangelidi O, Carr SB, et al.; CF-EpinNet collaboration. Data resource profile: the UK Cystic Fibrosis Registry. *Int J Epidemiol.* 2018;47:9–10e.
21. Quanjer PH, Stanojevic S, Cole TJ, et al.; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J.* 2012;40:1324–1343.
22. Taylor-Robinson D, Whitehead M, Diderichsen F, et al. Understanding the natural progression in %FEV1 decline in patients with cystic fibrosis: a longitudinal study. *Thorax.* 2012;67:860–866.
23. Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol.* 2001;153:345–352.
24. Konstan MW, Morgan WJ, Butler SM, Pasta DJ, Craib ML, Silva SJ, et al. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *J Pediatr.* 2007;151:134–139, 139.e1.
25. Konstan MW, Wagener JS, Vandevanter DR, et al. Risk factors for rate of decline in FEV1 in adults with cystic fibrosis. *J Cyst Fibros.* 2012;11:405–411.
26. Vandenbranden SL, McMullen A, Schechter MS, et al.; Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Lung function decline from adolescence to young adulthood in cystic fibrosis. *Pediatr Pulmonol.* 2012;47:135–143.
27. Schechter MS. Non-genetic influences on cystic fibrosis lung disease: the role of sociodemographic characteristics, environmental exposures, and healthcare interventions. *Semin Respir Crit Care Med.* 2003;24:639–652.
28. 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;161:233–241.
29. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 1986;51:1173–1182.
30. Schluchter MD, Piccorelli AV. Shared parameter models for joint analysis of longitudinal and survival data with left truncation due to delayed entry – applications to cystic fibrosis. *Stat Methods Med Res.* 2019;28:1489–1507.