

Validation and use of a parametric model for projecting cystic fibrosis survivorship beyond observed data: a birth cohort analysis

Abaigeal D Jackson,^{1,2} Leslie Daly,¹ Andrew L Jackson,³ Cecily Kelleher,¹ Bruce C Marshall,⁴ Hebe B Quinton,⁵ Godfrey Fletcher,² Mary Harrington,² Shijun Zhou,² Edward F McKone,⁶ Charles Gallagher,⁶ Linda Foley,² Patricia Fitzpatrick¹

¹UCD School of Public Health, Physiotherapy and Population Science, Woodview House, University College Dublin, Belfield, Dublin, Ireland

²The Cystic Fibrosis Registry of Ireland, Woodview House, University College Dublin, Belfield, Dublin, Ireland

³Department of Zoology, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland

⁴Cystic Fibrosis Foundation, Bethesda, Maryland, USA

⁵Department of Medicine, Dartmouth Medical School, Hanover, New Hampshire, USA

⁶St. Vincent's University Hospital, Elm Park, Dublin, Ireland

Correspondence to

Abaigeal Jackson, UCD School of Public Health, Physiotherapy and Population Science, Woodview House, University College Dublin, Belfield, Dublin 4, Ireland; abaigeal.jackson@ucd.ie

In memoriam of Mrs Linda Foley, Chief Executive of the Cystic Fibrosis Registry of Ireland (2001–2009).

Received 21 July 2010
Accepted 19 March 2011
Published Online First
8 June 2011

ABSTRACT

Background The current lifetable approach to survival estimation is favoured by CF registries. Recognising the limitation of this approach, we examined the utility of a parametric survival model to project birth cohort survival estimates beyond the follow-up period, where short duration of follow-up meant median survival estimates were indeterminable.

Methods Parametric models were fitted to observed survivorship data from the US CF Foundation (CFF) Patient Registry 1980–1994 birth cohort. Model-predicted median survival was estimated. The best fitting model was applied to a Cystic Fibrosis Registry of Ireland dataset to allow an evaluation of the model's ability to estimate predicted median survival. This involved a comparison of birth cohort lifetable predicted and observed (Kaplan–Meier) median survival estimates.

Results A Weibull model with main effects of gender and birth cohort was developed using a US CFF dataset (n=13 115) for which median survival was not directly estimable. Birth cohort lifetable predicted median survival for male and female patients born between 1985 and 1994 and surviving their first birthday was 50.9 and 42.4 years respectively. To evaluate the accuracy of a Weibull model in predicting median survival, a model was developed for the 1980–1984 Cystic Fibrosis Registry of Ireland birth cohort (n=243), which had an observed (Kaplan–Meier) median survival of 27.7 years. Model-predicted median survival estimates were calculated using data censored at different follow-up periods. The estimates converged to the true value as length of follow-up increased.

Conclusions Accurate prognostic information that is clinically critical for care of patients affected by rare, life-limiting disorders can be provided by parametric survival models. Problems associated with short duration of follow-up for recent birth cohorts can be overcome using this approach, providing better opportunities to monitor survival and plan services locally.

INTRODUCTION

Cystic fibrosis (CF) registries are the principal source of data for analysis of CF survival.¹ The US CF Registry is one of the largest and longest established, reporting data from over 25 000 patients receiving care at CF Foundation (CFF) accredited centres since 1982. Median survival is

Key messages

What is the key question?

Can median survival estimates be projected beyond observed survivorship data for recent birth cohorts of people with cystic fibrosis (CF) when more than half of the cohort is still alive?

What is the bottom line?

Parametric modelling of CF birth cohort data can overcome the challenge of estimating birth cohort median survival following a short follow-up period.

Why read on?

Accurate prognostic information can be provided by parametric models of CF birth cohort data, providing better opportunities to monitor and plan CF services.

a common measure used by registries to monitor life expectancy for patients with CF. Median survival estimation follows an entire CF population until half are deceased, and is the age at which half of the population have died. Median survival can be derived using two different approaches: current (period) and cohort lifetables.^{2–8} However, the application of each has limitations in CF⁹ and has resulted in failure to adopt a simple standardised approach internationally.

The current lifetable approach takes age-specific mortality rates observed in a calendar year and estimates a 'current lifetable predicted median survival' for a hypothetical population by assuming those mortality rates validly estimate future rates and remain constant over the population's lifetime.¹⁰ However mortality rates are not static; notable improvements have been observed during childhood in recent years.^{9 11}

The alternative method uses a birth cohort lifetable to estimate 'birth cohort lifetable observed median survival' and requires all CF births in a given period to be identified and followed until death or until more than 50% are deceased for estimation of median survival. Cohorts of recent births will not have an estimate for many years because more than half are still alive.¹²

Because of the inability of many CF registries to observe birth cohort lifetable median survival for



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://thorax.bmj.com/site/about/unlocked.xhtml>

cohorts of recent births, the current lifetable approach is more frequently applied to monitor temporal trends in median survival.^{3,4} For instance, in 2008, the US CFF Patient Registry reported that 46.3% of patients with CF were aged ≥ 18 years,³ and that the current lifetable predicted median survival estimate was 37.4 years.

We examined the utility of a parametric survival model to obtain estimates of birth cohort lifetable predicted median survival beyond the follow-up period, where short duration of follow-up by CF registries meant birth cohort lifetable observed median survival estimates were otherwise indeterminable.

STUDY DATA

This analysis is based on CF data from two patient registries. The first dataset was derived from the US CFF Patient Registry. This registry confirms deaths and death dates by direct contact with the CF centres. US death certification records were unavailable and patients lost to follow-up in 2007 (did not attend a CFF clinic in that year) were therefore excluded from this analysis. The second dataset was defined using data from the Cystic Fibrosis Registry of Ireland (CFRI) and a listing of deaths compiled from three sources: Central Statistics Office registered deaths when CF was reported as the underlying cause (1980–2007), CF centre records of CF deaths (2002–2007), and CF patient association recorded deaths (1986–2007). Study patients not reported as deceased were therefore presumed to be alive.

Each study population comprised a birth cohort of patients with CF born between 1980 and 1994 and diagnosed by 31 December 2007 ('study end date'). Patients who underwent solid-organ transplantation were not excluded or censored at transplantation to ensure consistency across datasets. The following data were provided for each patient: year of birth, gender, vital status on study end date, and survival time from birth to death or study end date. We calculated survival for all study patients from date of birth, although left truncation at either date of diagnosis or entry into the CF registry would be the most appropriate treatment of the data. Left truncation was precluded by the absence of date of diagnosis data for Ireland decedents not enrolled on the CFRI (ie, study patients identified from registered death records). Although the US and Irish CF registries report the highest CF population ascertainment rates internationally (approximately 90%), median survival estimates may be unavoidably biased because all people with CF born in a specific time period may not have been detected and followed over the period.¹³

Table 1 shows differences between CFF and CFRI birth cohorts. The number of US registry-enrolled patients with CF born between 1980 and 1994 is 20 times that of the Irish registry, and the number of birth cohort deaths by 2007 is nearly twice as many for CFRI (30.2%) compared with CFF (17.4%).

Deaths in CFF and CFRI birth cohorts were then examined using three, 5-year birth cohorts (1980–1984, 1985–1989, 1990–1994) (table 2). Such a small proportion of deaths occurred during the follow-up period (13 years) of the 1990–1994 birth cohort that it was deemed appropriate to combine the 1985–1989 and 1990–1994 birth cohorts. Subsequently, birth cohort life-tables were calculated and survivorship curves generated for each country, for one 5-year and one 10-year birth cohort (1980–1984, 1985–1994) using the Kaplan–Meier method (SURV and SURVFIT functions) in R software version 2.10.1 (downloadable from <http://www.r-project.org/>).¹⁴ Although less than 50% of the CFRI 1980–1984 birth cohort

Table 1 Characteristics of patients born between 1980 and 1994 enrolled on the US Cystic Fibrosis Foundation Patient Registry and the Cystic Fibrosis Registry of Ireland, with known vital status, and who were diagnosed with cystic fibrosis by 2007

Characteristics	US Cystic Fibrosis Foundation Patient Registry	Cystic Fibrosis Registry of Ireland
Patients with CF, 1980–1994 birth cohorts	13 164	659
Gender: male	6730 (51.1%)	354 (53.7%)
Birth cohort deaths by 2007	2290	199
Deaths in under 1 year olds	50	35
Person-years of follow-up	248 995.6	11 728.0
Deaths per 10 000 person-years	92.0	169.7

had died by 2007, the observed median survival could be determined using the Kaplan–Meier method. The point in time to which 50% of the cohort survives is not necessarily the same as the median for the Kaplan–Meier derived cumulative survival function. This can occur when a lifetable contains censored observations prior to the 50th percentile of the cumulative survival function.

ANALYSIS STRATEGY

Various parametric models were used to fit survival curves for birth cohort lifetable predicted median survival estimation. Because of its large size, the pooled US CFF birth cohort dataset (1980–1994) was used to select the best fitting parametric survival model. Model parameters were then estimated for the US dataset using the variables gender and birth cohort (1980–1984, 1985–1994). The model structure that best fit the US dataset was applied to the CFRI dataset in order to evaluate the ability of the model to estimate birth cohort lifetable predicted median survival. Model parameters were calculated independently for the CFRI dataset.

INTRODUCTION TO MODEL DEVELOPMENT

A parametric model assumes that observed data follows a particular distribution (eg, exponential, Weibull) and that one of those different distributions will best fit CF survival data. Unlike non-parametric models (eg, Cox proportional hazards), extrapolation of survival estimates beyond the study observation period is possible using a parametric approach.¹⁵ However, the benefit of projecting a survival estimate beyond observed data is tempered by the sensitivity of the estimate (and its associated degrees of uncertainty) to the underlying model structure and is governed by the size and duration of follow-up of the observed birth cohort.

Patients with CF who died before their first birthday were excluded during model development because high mortality

Table 2 Deaths observed by 2007 for three 5-year birth cohorts (1980–1984, 1985–1989, 1990–1994) of patients enrolled on the US Cystic Fibrosis Foundation Patient Registry and the Cystic Fibrosis Registry of Ireland

	US Cystic Fibrosis Foundation Patient Registry		Cystic Fibrosis Registry of Ireland	
	Birth cohort (N)	Deaths (%)	Birth cohort (N)	Deaths (%)
1980–1984	4070	32	261	47.9
1985–1989	4510	15.6	228	21.5
1990–1994	4584	6.2	170	14.7

rates during the first year of life led to poor model fit. Additionally, this action unintentionally benefitted the analysis because potential bias introduced by the unavoidable omission of unrecognised infant deaths from CFF³ 16 and CFRI datasets was avoided. The parametric model and the corresponding birth cohort life table predicted median survival estimates are for patients with CF surviving their first birthday.

Model parameter estimation, which was performed independently for CFF and CFRI datasets, involved adding gender and birth cohort (1980–1984, 1985–1994) as potential explanatory variables to the model. Differences between the two cohorts and between male and female survival estimates were examined (table 3). Fully separated models (stratified by both gender and birth cohort) were excluded from the analysis a priori because of the small proportion of deaths observed in US male patients born between 1985 and 1994 (9%) and US female patients born between 1985 and 1994 (12%). Parametric model fitting and parameter estimation was performed using the SURVREG function in R.

MODEL SELECTION AND FITTING USING US CF FOUNDATION PATIENT REGISTRY DATA

Four model structures, each with a distinct underlying survivorship function for fitting observed data to parametric models, were deemed suitable for examination: exponential, Weibull, log-logistic and log-normal. Initially, all US CFF data were entered into the model, pooling observations across birth cohorts (1980–1994) and gender to help decide which model structure was most appropriate for the data (table 3, model A).

Akaike’s Information Criterion (AIC) was used to determine which underlying model best fitted observed survivorship, with lower values indicating a better fit. Because log-likelihood is an additive property of a model and dataset,¹⁷ AIC values can be summed across non-overlapping stratifications of the same dataset to yield an overall AIC value for the entire dataset.¹⁸ Birth cohort life table predicted median survival and 95% CIs were estimated using the SURVREG function reverse prediction method.¹⁴

Using the pooled US dataset (1980–1994) of patients with CF surviving their first birthday (13 115 observations), the Weibull (AIC =24 060.3) model best explained the data (figure 1, table 3). On examining differences in survivorship between gender and birth cohorts, the best model contained gender and birth cohort (1980–1984, 1985–1994) as the main effects under a Weibull

Table 3 Akaike Information Criteria (AIC) for Weibull models with explanatory variables of gender and birth cohort (1980–1984 and 1985–1994) for patients enrolled on the US Cystic Fibrosis Foundation Patient Registry surviving their first birthday

Model	Explanatory variable(s)	Weibull model AIC
A	Null (no variables)	24 302.8
B	Gender	24 232.1
C	Birth cohort	24 128.7
D	Gender and birth cohort as main effects	24 060.3
E	Gender and birth cohort with interaction	24 062.2
F	Separated models for gender	24 232.4
	Male	10 828.4
	Female	13 404.0
G	Separated models for birth cohort	24 108.3
	1980–1984	13 000.1
	1985–1994	11 108.2

distribution (table 3, model D), and was marginally better than the model containing gender and birth cohort with an interaction effect (model E).

Under the model with main effects of gender and birth cohort, US male patients were shown to live longer than female patients and survival improved between the 1980–1984 and 1985–1994 birth cohorts (figure 2A, table 4). Birth cohort life table predicted median survival for patients surviving their first birthday was 37.8 and 31.5 years for male and female patients born between 1980 and 1984, and 50.9 years and 42.4 years for male and female patients born between 1985 and 1994.

PARAMETER FITTING USING CFRI DATA

Because the Weibull model was deemed to best fit US CFF survivorship data, the model was applied to the CFRI dataset. Estimation of CFRI model parameters took place independently of the CFF model parameter estimation. A dataset of 624 CFRI patients with CF born between 1980 and 1994 and surviving their first birthday was entered into the model. An AIC of 1659 was determined for the null Weibull model (ie, no explanatory factors were used in the model). Inclusion of explanatory factors improved model fit and a Weibull model containing gender and birth cohort as the main effects was the best fitting (AIC=1633).

Birth cohort life table predicted median survival estimates for male and female patients surviving their first birthday were 32.3 and 24.7 years respectively for those born between 1980 and 1984, and 51.1 and 39.0 years for those born between 1985 and 1994 (figure 2B, table 5). Predicted median survival estimates increased between the birth cohorts by 18.8 years for male patients and 14.3 years for female patients.

COMPARISON OF CFRI BIRTH COHORT LIFE TABLE OBSERVED (KAPLAN–MEIER) AND PREDICTED MEDIAN SURVIVAL ESTIMATES

While birth cohort life table predicted median survival estimates were calculated for both CFF and CFRI study populations, the

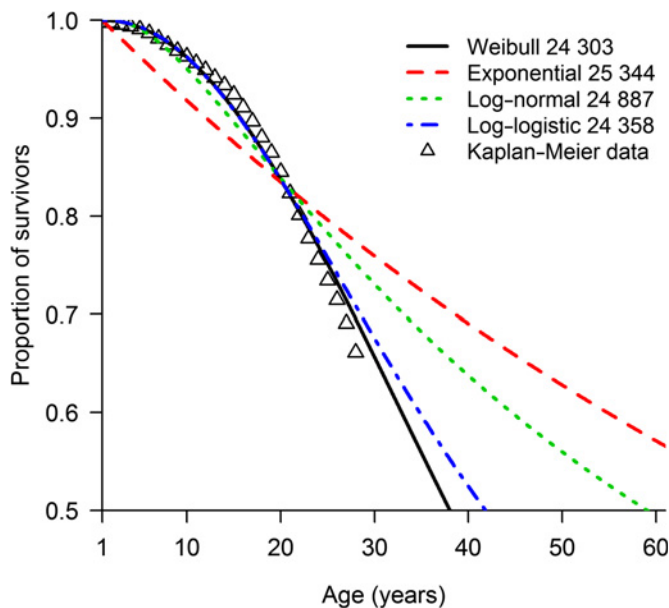
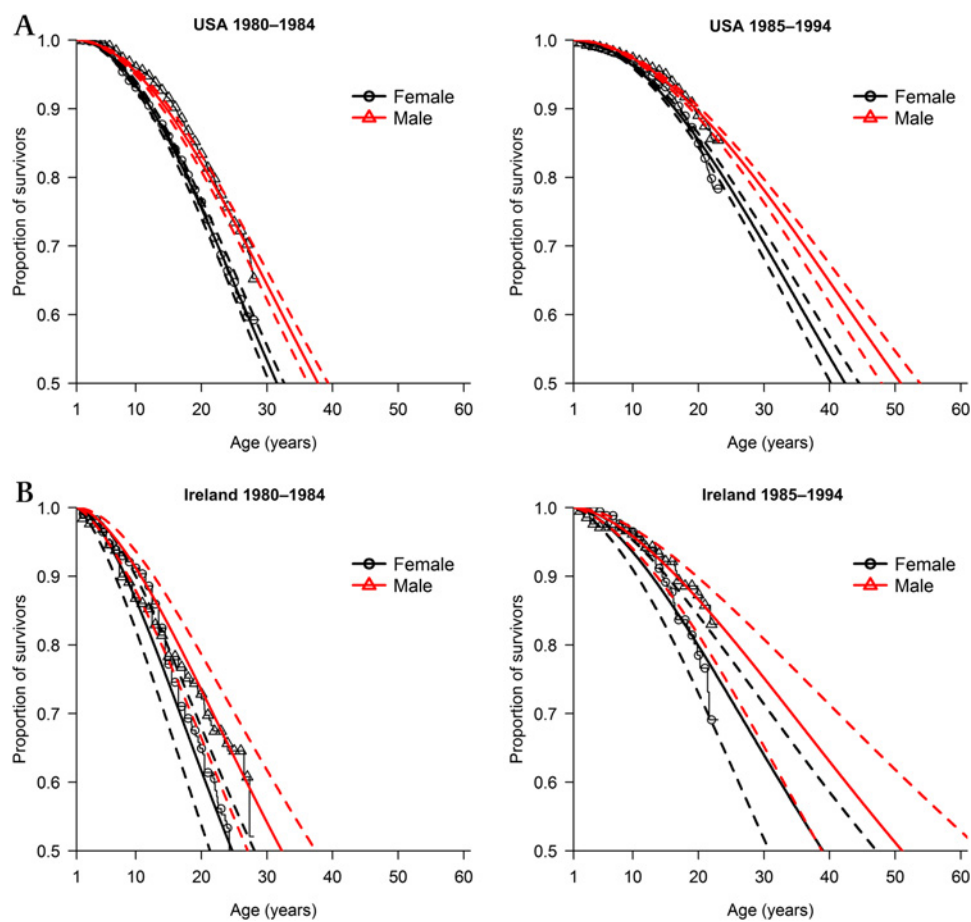


Figure 1 Kaplan–Meier survivorship curve and parametric models with associated Akaike Information Criteria for patients born between 1980 and 1994 enrolled on the US Cystic Fibrosis Foundation Patient Registry who survived their first birthday.

Figure 2 Kaplan–Meier survivorship curves (females = ○, males = △) with fitted Weibull model with main effects of gender (female = black line, male = red line) and birth cohort and corresponding 95% CIs (dashed lines) for patients with cystic fibrosis surviving their first birthday and enrolled on the US Cystic Fibrosis Foundation Patient Registry (A) and the Cystic Fibrosis Registry of Ireland (B).



latter dataset allowed an evaluation of the ability of the parametric model to predict birth cohort lifetable observed median survival. A birth cohort lifetable observed median survival estimate was derived for CFRI patients with CF born between 1980 and 1984 and surviving their first birthday using the Kaplan–Meier method (see Study data). Model-estimated birth cohort lifetable predicted median survival was compared with the actual observed (Kaplan–Meier) median survival to determine whether the selected model was capable of reliably estimating predicted median survival. A full evaluation of the suitability of the model for estimating survival will only be possible once a dataset becomes available which has followed all patients in a birth cohort until death.

As the actual (Kaplan–Meier) median survival of the CFRI 1980–1984 birth cohort was observed for male patients, but not female patients, the comparison of model-predicted and observed median survival estimates required male and female birth cohort datasets to be pooled. A Weibull model was applied

to the CFRI 1980–1984 birth cohort consisting of 243 patients with CF surviving their first birthday (cf, table 3, model G: 1980–1984). The performance of the model in predicting median survival after short duration of follow-up was examined by censoring observed survivorship data (figure 3). Initially, survivorship data were censored after 5 years; therefore only follow-up data to the 31 December 1989 were used in the model. A birth cohort lifetable predicted median survival of 40.0 years was estimated based on only 5 years of follow-up.

The follow-up period was then extended a year at a time until 2007. A new model was developed each time and model-predicted median survival was re-estimated on each occasion. Convergence of the birth cohort lifetable predicted to the observed median survival estimate (for patients with CF surviving their first birthday) of 27.7 years (lower 95% CI: 24.8, upper 95% CI: indeterminate) was good, particularly from 1995 onwards, which corresponds to 20% observed mortality in the 1980–1984 birth cohort.

Table 4 Birth cohort lifetable predicted median survival estimates and 95% CIs for a Weibull model with main effects of gender and birth cohort for patients enrolled on the US Cystic Fibrosis Foundation Patient Registry surviving their first birthday

Gender and birth cohort	Median survival estimate in years (95% CI)
Male births 1980–1984	37.8 (36.2 to 39.3)
Female births 1980–1984	31.5 (30.4 to 32.6)
Male births 1985–1994	50.9 (48.0 to 53.8)
Female births 1985–1994	42.4 (40.2 to 44.5)

Table 5 Birth cohort lifetable predicted median survival estimates and 95% CIs for a Weibull model with main effects of gender and birth cohort for patients enrolled on the Cystic Fibrosis Registry of Ireland surviving their first birthday

Gender and birth cohort	Median survival estimate in years (95% CI)
Male births 1980–1984	32.3 (27.1 to 37.5)
Female births 1980–1984	24.7 (21.3 to 28.1)
Male births 1985–1994	51.1 (38.8 to 63.3)
Female births 1985–1994	39.0 (30.8 to 47.1)

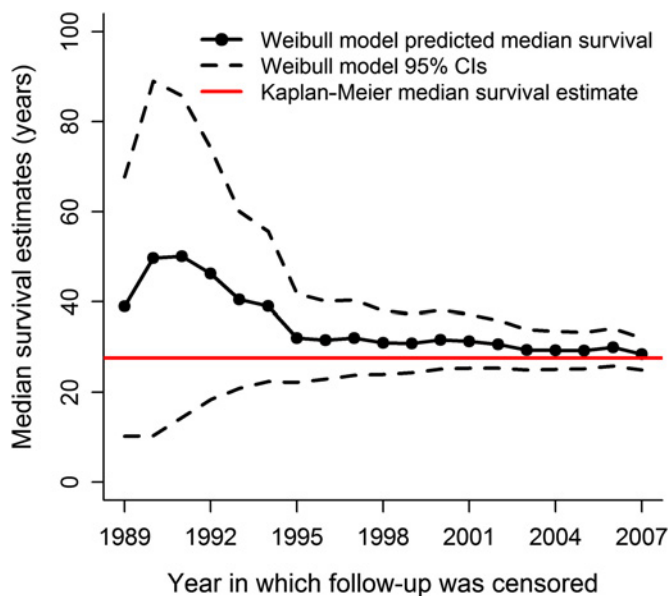


Figure 3 Median predicted survival and 95% CIs for patients born between 1980 and 1984 enrolled on the Cystic Fibrosis Registry of Ireland and surviving their first birthday, estimated under a Weibull model during a 23-year follow-up period, in which follow-up was censored in each calendar year. The actual (Kaplan–Meier) median survival for patients with cystic fibrosis surviving their first birthday observed at the study end point (2007) is also shown.

DISCUSSION

By developing parametric models of CF birth cohort data to project survivorship, we have started to overcome the challenge of median survival estimation associated with short duration of follow-up of recent cohorts. This approach is suited to survival estimation for disorders where better treatment has resulted in continued improvement in survivorship. Although limitations remain when estimating survival in rare disease populations, regardless of methodological approach,⁹ we have shown that our model can be applied to registry datasets of variable size, demonstrating its utility as a tool to estimate survival at a local level.

Better treatment has resulted in continued improvement in CF survival, particularly in childhood, yet the current life table method, which applies static age-specific mortality rates to a hypothetical population in order to project estimates of life expectancy, is favoured by CF registries. At best, annual median predicted CF survival estimates derived by current life tables can only provide a snapshot of a moving target.¹³ For conditions with fluctuating age-specific mortality rates, longitudinal techniques, that is, birth cohort life tables, may be more appropriate for monitoring survival. However, prospective follow-up of birth cohorts by many CF registries has not yet yielded an observed median survival. Birth cohort follow-up can be undertaken retrospectively and although limitations to this approach exist, using registered death information to supplement registry data reduces the probability of positively biasing survival estimates, and allows parametric modelling to proceed.

One advantage of predicting birth cohort median survival is that median survival can be estimated by gender and year of birth. Therefore, counselling parents of newly diagnosed patients on the prospect of their child's survival can be appropriately tailored. This is an improvement on current life table estimates, which are typically reported for an entire registry

population. For example, the US CFF reported a current life table median survival estimate of 37.4 years in 2008,³ whereas model-derived birth cohort median predicted survival estimates suggest that US male and female patients born between 1985 and 1994 and surviving their first birthday live a median of 50.9 and 42.4 years respectively. As in others studies of CF survival which considered gender and birth cohort as explanatory factors, this study found that the period effect was the strongest determinant of survival into adulthood.¹⁹ Second, clinical management of patients with CF is complex and the survival benefit from novel therapies, early diagnosis and specialist, multi-disciplinary care remain unclear. Projecting birth cohort survivorship beyond the treatment period using a parametric approach could provide better insights into the effectiveness of current clinical practices.

A comparison of model estimates of median predicted survival between the 1980–1984 and 1985–1994 birth cohorts reaffirms that CF survival is improving in both the USA and Ireland. As in other studies, a gender gap in CF survival can be observed,^{20 21} however data comparing observed survival by gender across sequential birth cohorts are not often published,^{5 22 23} making it difficult to interpret our findings.

Differences in median survival estimates appear to exist internationally in this study and other studies.²³ In addition to factors such as variability in diagnostic practices, genotype distribution and patient and transplantation management strategies between countries,^{24–26} the methodological approach used to estimate survival should be considered as a possible source of discordance during international survival estimate comparison. Cross-sectional comparison (ie, at a given time point) of international median predicted survival estimates derived using current life tables has been shown to be misleading.²⁷ Comparison of median age at death estimates derived from decedent CF populations has also been undertaken,²⁴ but this measure of survival is considered unsatisfactory because it underestimates the observed median survival derived from following an entire CF population.²⁷ Although this study has shown that parametric models can be fitted to registry datasets to derive estimates of median predicted survival, registry information has inherent limitations that can affect the reliability of model-derived survival estimates.

CF registries acknowledge that their datasets are not entirely representative of the CF population, and report population coverage levels of approximately 90%.^{20 28} Because patient consent is required for registry enrolment, registries are often unable to observe those who died before being invited to enrol, those opting not to participate, and others yet to be identified by the registry. To avoid the introduction of bias into birth cohort life tables, all affected people need to be included. Differences in median survival estimates may be partly explained by potential biases in data collection by CF registries. Unlike the US datasets, national mortality records were included in the Irish birth cohort datasets, although deceased patients with CF in Ireland may not have been identified if CF was not reported as the underlying cause of death (eg, trauma, suicide), but rather a factor contributing to death. The omission of unrecognised deceased patients with CF from the US CFF Patient Registry prior to 1986, particularly childhood mortality data, which are not insubstantial,¹⁶ may have resulted in US survival estimates being positively biased. Although the large size of the US datasets allows more precise estimates of birth cohort life table predicted median survival to be derived, the effect of bias on survival estimation is not reduced.

Although the inclusion of additional explanatory variables (potentially available for prospectively followed birth cohorts)

could further refine the predictive ability of this model, this was not possible for the analysis of the Irish data. Instead, our aim was to examine whether survival estimates could be projected beyond observed follow-up data in the first instance and identify a suitable model structure which could potentially be adopted by others.

Estimates of survival are important for planning service requirements and evaluating interventions, but in reality, how long any newborn baby with CF will live cannot be known. As with any model projecting survival estimates beyond observed data, population ascertainment, size and duration of follow-up of observed data affect the sensitivity of our model. When observed and predicted birth cohort life table median survival estimates were examined for our validation cohort, we found that estimates were comparable but also that the model was valid in a population smaller than that used to develop the model.

In conclusion, parametric models of birth cohorts can overcome the challenge of estimating birth cohort life table predicted median survival following a short follow-up period for a disorder experiencing continued improvement in survivorship. This approach may provide better opportunities to monitor survival and plan CF services at the local level.

Acknowledgements We thank the UK and Australian CF registries for their willingness to contribute data to this analysis, but who were restricted in doing so by the absence of registry recorded deaths from their registries prior to 1998. We also are grateful to Professor Philip Farrell and Professor Chris Goss for critically reviewing an earlier draft.

Funding This work was supported by the Cystic Fibrosis Registry of Ireland Executive Council. Health Research Board, Ireland (RP/2007/249).

Competing interests None declared.

Ethics approval Cystic fibrosis registry data were used in this study, and the registry was given ethical approval by each Ireland cystic fibrosis centre and clinic to collect data on cystic fibrosis patients.

Contributors All authors approved the final submitted version. Authors' specific contributions are listed below: Dr Abaigeal D Jackson: study concept and design, data acquisition, data analysis and interpretation, manuscript preparation. Professor Leslie Daly: data analysis and interpretation, study design, study supervision. Dr Andrew L Jackson: data analysis and interpretation, review of manuscript drafts. Professor Cecily Kelleher: study concept and design, review of manuscript drafts, study supervision. Dr Bruce C Marshall: data interpretation, critical revision. Dr Hebe B Quinton: data acquisition, interpretation, critical revision. Mr Godfrey Fletcher: study concept and design, critical revision. Ms Mary Harrington: data acquisition, review of manuscript drafts. Mr Shijun Zhou: data acquisition, review of manuscript drafts. Dr Edward F McKone: review of manuscript drafts, data interpretation. Dr Charles Gallagher: review of manuscript drafts, data interpretation. Mrs Linda Foley: study concept and design, study supervision. Dr Patricia Fitzpatrick: data analysis, study supervision.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **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**:229–37.
2. **Bellis G**, Cazes MH, Parant A, *et al*. Cystic fibrosis mortality trends in France. *J Cyst Fibros* 2007;**6**:179–86.
3. **Cystic Fibrosis Foundation**. *Cystic Fibrosis Foundation Patient Registry annual data report 2008*. Bethesda, MD: 2009.
4. **Cystic Fibrosis Trust**. *UK CF Registry annual data report 2008*. Bromley, Kent: 2009.
5. **Dodge JA**, Lewis PA, Stanton M, *et al*. Cystic fibrosis mortality and survival in the UK: 1947–2003. *Eur Respir J* 2007;**29**:522–6.
6. **Frederiksen B**, Lannig S, Koch C, *et al*. Improved survival in the Danish center-treated cystic fibrosis patients: results of aggressive treatment. *Pediatr Pulmonol* 1996;**21**:153–8.
7. **Lai HJ**, Cheng Y, Farrell PM. The survival advantage of patients with cystic fibrosis diagnosed through neonatal screening: evidence from the United States Cystic Fibrosis Foundation Registry data. *J Pediatr* 2005;**147**(Suppl 3):S57–63.
8. **Stern M**, Wiedemann B, Wenzlaff P, *et al*. From registry to quality management: the German Cystic Fibrosis Quality Assessment Project, 1995–2006. *Eur Respir J* 2008;**31**:29–35.
9. **Kulich M**, Rosenfeld M, Goss CH, *et al*. Improved survival among young patients with cystic fibrosis. *J Pediatr* 2003;**142**:631–7.
10. **Daly LE**, Bourke GJ. *Interpretation and uses of medical statistics*. London: Blackwell Science, 2000.
11. **Dodge JA**, Lewis PA. Cystic fibrosis is no longer an important cause of childhood death in the UK. *Arch Dis Child* 2005;**90**:547.
12. **Lewis PA**. Inferences for health provision from survival data in cystic fibrosis. *Arch Dis Child* 1998;**79**:297–9.
13. **Corey M**. Survival estimates in cystic fibrosis: snapshots of a moving target. *Pediatr Pulmonol* 1996;**21**:2.
14. **R Development Core Team**. *R Software Version 2.10.1*. Vienna, Austria: R Foundation for Statistical Computing, 2008.
15. **Crawley MJ**. *Statistics: an introduction using R*. West Sussex, England: Wiley & Sons Ltd, 2007.
16. **Nick JA**, Chacon CS, Brayshaw SJ, *et al*. Effects of gender and age of diagnosis on disease progression in long-term cystic fibrosis survivors. *Am J Respir Crit Care Med* 2010;**182**:614–26.
17. **Edwards AWF**. *Likelihood*. Baltimore: Johns Hopkins University Press, 1992.
18. **Burnham KP**, Anderson DR. *Model selection and multimodal inference: a practical information-theoretic approach*. New York: Springer, 2002.
19. **Assael BM**, Castellani C, Ocampo MB, *et al*. Epidemiology and survival analysis of cystic fibrosis in an area of intense neonatal screening over 30 years. *Am J Epidemiol* 2002;**156**:397–401.
20. **FitzSimmons SC**. The changing epidemiology of cystic fibrosis. *J Pediatr* 1993;**122**:1–9.
21. **Rosenfeld M**, Davis R, FitzSimmons S, *et al*. Gender gap in cystic fibrosis mortality. *Am J Epidemiol* 1997;**145**:794–803.
22. **Dodge JA**, Morison S, Lewis PA, *et al*. Incidence, population, and survival of cystic fibrosis in the UK, 1968–95. *Arch Dis Child* 1997;**77**:493–6.
23. **Slieker MG**, Uiterwaal CS, Sinaasappel M, *et al*. Birth prevalence and survival in cystic fibrosis: a national cohort study in the Netherlands. *Chest* 2005;**128**:2309–15.
24. **Fogarty A**, Hubbard R, Britton J. International comparison of median age at death from cystic fibrosis. *Chest* 2000;**117**:5.
25. **McCormick J**, Mehta M, Olesen HV, *et al*. Comparative demographics of the European cystic fibrosis population: a cross-sectional database analysis. *Lancet* 2010;**375**:1007–13.
26. **McKone EF**, Emerson SS, Edwards KL, *et al*. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet* 2003;**361**:1671–6.
27. **Jackson AD**, Daly L, Kelleher C, *et al*. The application of current life table methods to compare median survival internationally is limited. *J Cyst Fibros* 2011;**10**:62–5.
28. **Cystic Fibrosis Registry of Ireland**. *Annual report 2007*. Dublin, 2008.