LONGITUDINAL AND CROSS-SECTIONAL MODELLING OF HEALTH RELATED QUALITY OF LIFE IN PEOPLE WITH CYSTIC FIBROSIS

Margaret Anne Hurley[1]

College of Health and Wellbeing, University of Central Lancashire, Preston, PR1 2HE, UK

Janice Abbott

School of Psychology, University of Central Lancashire, Preston, PR1 2HE, UK

Alison M. Morton, Steven P. Conway

Adult Cystic Fibrosis Unit, St James' Hospital, Leeds, LS9 7TF, UK

Abstract. *People with cystic fibrosis (CF) must endure up to four hours treatment per day to maintain health and are vulnerable to complications. The Cystic Fibrosis Quality of Life Questionnaire was developed to measure health related quality of life (HRQoL) in the UK. Most studies on HRQoL are cross-sectional in design with HRQoL measured once per patient. However, the Cystic Fibrosis Quality of Life Questionnaire has been used to monitor HRQoL longitudinally with measures taken over a 12 year period at one clinic in the UK. These data were modelled with a binomial distribution for a domain score and with fixed and random coefficients for the patient-level clinical and demographic variables. The longitudinal study included 182 patients whose HRQoL was first measured within a single calendar year and were then followed-up. These data provided an opportunity to compare, directly and by simulation, the modelling of a cross-sectional with a longitudinal study and so provided insights into the statistical merits of longitudinal studies compared to cross-sectional studies in HRQoL.*

Keywords. *Binomial distribution, CFQoL, random effect, HRQoL, logistic regression.*

1. INTRODUCTION

In the white population, cystic fibrosis (CF) is the most common autosomal recessive condition (Rosenstein, 1998). For people with this condition, the lungs, digestive system and other organs become affected by thick and sticky mucus because their cells are unable to regulate the production of sodium and chloride. In 2013 in the UK alone, 10,000 children and adults were receiving care in CF clinics

[1] MAHurley@uclan.ac.uk

and were registered with the UK CF Registry (UK Cystic Fibrosis Registry, 2014). People with CF require a treatment regimen of up to four hours per day to maintain optimal health and are vulnerable to many complications such as CF related diabetes, pneumothorax, osteoporosis and lung infections such as with *B. cepacia* complex. They also find it a significant challenge to obtain sufficient nutrition from their dietary intake. Despite this, in recent years in the UK at least, CF has ceased to be a disease only of childhood since now, with good medical care and rigorous daily treatment, many people with CF reach adulthood and move on into adult life, enter training or attend university, obtain employment and participate in adult social and family life (Laborde-Castérot et al., 2012). Even so, median age at death was 29 years for deaths occurring in 2013, though some individuals had survived into their late fifties and beyond (UK Cystic Fibrosis Registry, 2014). Clearly health related quality of life (HRQoL) is an important issue for people with CF and good psychological and mental well-being is essential, also, to allow people to comply with their treatment and maximise their longevity and their life chances (Abbot, 2009; Abbott et al., 2011).

Many cross-sectional studies have investigated the effects of clinical and demographic variables on HRQoL in CF (Gee et al., 2003, Gee et al., 2005, Wahl et al., 2005). In contrast there are few studies which have looked at HRQoL prospectively in a truly longitudinal manner (Abbott et al., 2013; Abbott et al., 2015). Such long-term studies may take a decade or more to complete (Abbott et al., 2013; Abbott et al., 2015) and research funding streams are rarely able to provide the necessary long-term support. Some studies which are described as longitudinal can be better understood as cross-sectional studies with either a single follow-up time for each patient (Sawicki et al., 2011) or have many assessments on the same patient but over a short time scale (Goldbeck et al., 2007) and it is questionable whether these are bona-fide longitudinal studies. So, in point of fact, real longitudinal studies in HRQoL are very rare whereas cross-sectional studies are not. From the point of view of epidemiology, longitudinal studies may be preferable to cross-sectional studies because they allow the assessment of cause and effect and longitudinal studies can show temporality which cross-sectional studies cannot always demonstrate. In the study of marketplace behaviours, the relative validity of longitudinal and cross-sectional surveys has been discussed in the context of common methods variance and valid causal inference (Rindfleisch et al., 2008). However, from the point of view of statistics, it is not clear what are the advantages of longitudinal over cross-sectional studies in terms of HRQoL model estimation and the assessment of the contribution of demographic and clinical changes to changes in HRQoL. This was due mostly to a lack of available genuine

longitudinal data in the past to inform a statistical comparison. Now, however, suitable longitudinal study data have become available (Abbott et al., 2013, 2015) and some comparisons have become feasible. This study provides a comparison based on direct modelling of the empirical data followed by a computer simulation.

2. THE LONGITUDINAL AND CROSS-SECTIONAL DATA

A longitudinal study engaged patients attending a single CF clinic in the north-east of England in 1998 and followed-up patients every two years until 2010. Thus patients had a maximum of 7 assessments in total. Patients sometimes failed to attend clinic and sometimes moved to a different clinic (for example when away at university) and sadly some patients died. Hence between 1 and 7 assessments were made for the patients and the first assessments were made during the calendar year 1998. At assessment, the demographic variables of gender and age and a range of clinical variables were recorded; weight, height, body mass index (BMI), lung function measured by forced expiratory volume in one second adjusted for age, gender and height (FEV_1% predicted), whether the patient was on oral nutritional supplements or was taking enteral tube feeds, whether the patient was diagnosed with CF-related diabetes, infected in the lung with *B. cepacia* complex, had a totally implantable vascular access device (TIVAD) fitted for antibiotic treatment, and whether they were on a waiting list for organ transplantation or had already received a transplantation.

HRQoL was measured using the CFQoL tool (Gee et al., 2000) which was developed and validated for the UK CF population and comprised nine domains of HRQoL; Physical functioning, Social functioning, Emotional responses, Treatment issues, Chest symptoms, Body image, Interpersonal relationships, Career concerns and Concerns for the future. The data for the demographic, clinical and HRQoL variables were complete for all patients when they attended clinic for assessment and no variable values were missing in the longitudinal data.

Table 1: Numbers of observations in the cross-sectional and the longitudinal data.

	Cross-sectional data		**Longitudinal data**	
Number of bi-annual times points	**Patients**	**Patient-assessments**	**Patients**	**Patient-assessments**
1	182	182	40	40
2			37	74
3			17	51
4			26	104
5			21	105
6			19	114
7			22	154
Total	**182**	**182**	**182**	**642**

The cross-sectional data were obtained by selecting the first assessment, taken during 1998, for each patient. If a longitudinal study had not gone ahead then the 1998 data would have comprised a cross-sectional study. Effectively the cross-sectional data was a subset of the longitudinal data collected over a single calendar year. Table 1 shows the number of patients (n=182) in the cross-sectional and longitudinal data and the numbers of patient-assessments in the longitudinal data (n=642 patient-assessments) together with the numbers of times the patients were assessed. Table 2 shows the summary statistics for the cross-sectional and the longitudinal data. Fortuitously, the cross-sectional and the longitudinal data have the same range for BMI, FEV_1% predicted and all HRQoL domains have a similar age range and the clinical variables are not very different between the two data sets. Had this not been the case, a comparison between cross-sectional and longitudinal models may not have been valid.

Table 2: Summary statistics of the demographic, clinical and health related quality of life measures in the cross-sectional and longitudinal data.

	Cross-sectional data 182 patients 182 patient-assessments		Longitudinal data 182 patients 642 patient-assessments	
	N (%)		N (%)	
Female gender	105 (57.7%)		379 (59.0%)	
CF related diabetes	56 (30.8%)		256 (39.9%)	
Oral supplements	60 (33.0%)		213 (33.2%)	
Enteral tube feeds	35 (19.2%)		104 (16.2%)	
TIVAD	86 (47.3%)		340 (53.0%)	
B. cepacia complex	14 (7.7%)		52 (8.1%)	
Listed for transplant	8 (4.4%)		26 (4.0%)	
Received transplant	17 (9.3%)		93 (14.5%)	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**
Age (years)	24.4 (6.9)	14-51	28.7 (8.1)	14-57
BMI	21.4 (2.9)	15.7-30.2	21.7 (2.8)	15.7-30.6
FEV_1% predicted	59.2 (24.2)	12-133	57.8 (23.2)	12-133
Physical functioning %	83.2 (21.4)	2-100	83.2 (20.2)	2-100
Social functioning %	83.4 (22.2)	0-100	84.4 (21.7)	0-100
Emotional responses %	79.1 (20.9)	8-100	79.3 (19.8)	8-100
Treatment issues %	78.0 (21.7)	0-100	75.4 (22.2)	0-100
Chest symptoms %	75.5 (24.1)	0-100	76.4 (22.8)	0-100
Body image %	68.8 (25.3)	0-100	69.9 (24.9)	0-100
Interpersonal relationships %	63.9 (22.1)	2-100	63.7 (22.4)	2-100
Career concerns %	63.5 (28.9)	0-100	61.5 (29.5)	0-100
Concerns for the future %	46.6 (24.6)	0-100	45.3 (24.7)	0-100

3. MODELS FOR THE CROSS-SECTIONAL AND LONGITUDINAL DATA

Using the CFQoL (Gee et al., 2000), the quality of life of a patient on a domain was measured as a percentage between 0% and 100% where the percentage was obtained as $100S/N$ where S was the patient's integer score for the domain and N was the maximum domain score achievable. The maximum varied between the nine domains; Physical functioning (50), Social functioning (20), Emotional responses (40), Treatment issues (15), Chest symptoms (20), Body image (15), Interpersonal relationships (50), Career concerns (20) and Concerns for the future (30). Modelling HRQoL presented challenges because of the well-understood 'ceiling' effect in that HRQoL is bounded above by 100% (Walters, 2009). For this reason a binomial regression model was selected with a logistic link function between the binomial probability and the linear predictor incorporating the demographic and clinical variables. The binomial probability (expressed as a percentage) was assumed to represent the patient's true HRQoL on the domain. The linear predictor regression model in turn was assumed to contain both fixed and random patient-specific regression coefficients in order to model the correlation between HRQoL measures on the same patient at the different times in the longitudinal data. The random regression coefficients were assumed to be normally distributed over the population of patients. The models were fitted using the software MLwiN v2.1 (Rasbash et al., 2009) and detailed descriptions of the algorithms used by MLwiN may be found in Goldstein (2011).

For the longitudinal data, exploratory analyses resulted in a final selected regression model in which the covariates FEV_1% predicted, BMI and the model intercept were included with random coefficients across patients, age was included with a fixed coefficient and all categorical variables had fixed coefficients of necessity. The choice between random or fixed coefficients for quantitative variable was based on an assessment of whether the variance of a candidate random term was negligible and also whether it had negligible covariance with other random coefficients. The cross-sectional data, with one assessment per patient, could not support the estimation of random coefficients for FEV_1% predicted or for BMI but could support the estimation of a random intercept. This was because the random coefficients for FEV_1% predicted and BMI were patient specific and one observation per patient in the cross-sectional data would prevent the estimation of the variance of the random coefficients. The random intercept however varied between patients and so the intercept variance was estimable from the cross-sectional data.

The binomial denominators for the nine HRQoL domains ranged from 15 to 50 and the comparison between using the cross-sectional data and using longitudinal data might depend on the size of the denominator; the smaller denominator possibly

producing less sensitive data resulting in less precise estimates than with a larger denominator. For this reason the comparison models were fitted to the CFQoL domain of Physical functioning with a denominator of 50 and the domain of Treatment issues with a denominator of 15. Tables 3 and 4 show the models fitted to the cross-sectional and longitudinal data for the domains of Physical functioning and Treatment issues respectively. Figure 1 and Figure 2 show the comparison bar charts of the model coefficients together with plus and minus one standard error.

Table 3: Comparison of the estimated model coefficients between the cross-sectional and the longitudinal data for the domain of Physical functioning.

	Cross-sectional data			Longitudinal data	
	Coefficient	**SE**		**Coefficient**	**SE**
Categorical Fixed			**Categorical Fixed**		
Female gender	0.05	0.26	Female gender	-0.12	0.21
CF related diabetes	0.25	0.27	CF related diabetes	0.03	0.09
Oral supplements	-0.04	0.29	Oral supplements	-0.05	0.09
Enteral tube feeds	-0.30	0.40	Enteral tube feeds	0.27	0.13
TIVAD	-0.65	0.27	TIVAD	-0.38	0.09
B. cepacia complex	-0.54	0.43	*B. cepacia* complex	-0.36	0.19
Listed for transplant	-0.93	0.62	Listed for transplant	-0.24	0.14
Received transplant	0.39	0.44	Received transplant	-0.36	0.22
Quantitative Fixed			**Quantitative Fixed**		
Age[1]	0.00	0.10	Age[1]	-0.11	0.03
			Quantitative Random		
BMI[2]	0.01	0.05	BMI[2] Mean	0.10	0.04
			BMI Variance	0.17	0.03
FEV_1% predicted[3]	-0.15	0.07	FEV_1% predicted[3] Mean	-0.20	0.05
Quantitative Random			FEV_1% predicted Variance	0.16	0.03
Intercept Mean	1.65	0.41	Intercept Mean	1.31	0.20
Intercept Variance	2.14	0.24	Intercept Variance	1.08	0.19

[1] 5-year increase, age centred on 30 years, [2] One unit decline, BMI centred on 20, [3] 10% decline, FEV_1% predicted centred on 30%.

Table 4: Comparison of model coefficients between the cross-sectional and the longitudinal data for the domain of Treatment issues.

	Cross-sectional data			Longitudinal data	
	Coefficient	**SE**		**Coefficient**	**SE**
Categorical Fixed			**Categorical Fixed**		
Female gender	-0.14	0.20	Female gender	0.02	0.16
CF related diabetes	0.12	0.22	CF related diabetes	-0.03	0.11
Oral supplements	-0.40	0.23	Oral supplements	-0.20	0.11
Enteral tube feeds	-0.05	0.31	Enteral tube feeds	0.06	0.17
TIVAD	-0.30	0.21	TIVAD	-0.40	0.10
B. cepacia complex	-0.35	0.33	*B. cepacia* complex	-0.43	0.20
Listed for transplant	-0.43	0.48	Listed for transplant	-0.45	0.17
Received transplant	0.18	0.35	Received transplant	0.63	0.22
Quantitative Fixed			**Quantitative Fixed**		
Age[1]	0.02	0.02	Age[1]	-0.20	0.04
			Quantitative Random		
BMI[2]	0.08	0.04	BMI[2] Mean	0.08	0.03
			BMI Variance	0.02	0.01
FEV_1% predicted[3]	-0.05	0.05	FEV_1% predicted[3] Mean	-0.05	0.05
Quantitative Random			FEV_1% predicted Variance	0.07	0.02
Intercept Mean	1.52	0.33	Intercept Mean	1.13	0.20
Intercept Variance	1.03	0.15	Intercept Variance	0.85	0.17

[1] 5-year increase, age centred on 30 years, [2]One unit decline, BMI centred on 20, [3]10% decline, FEV_1% predicted centred on 30%.

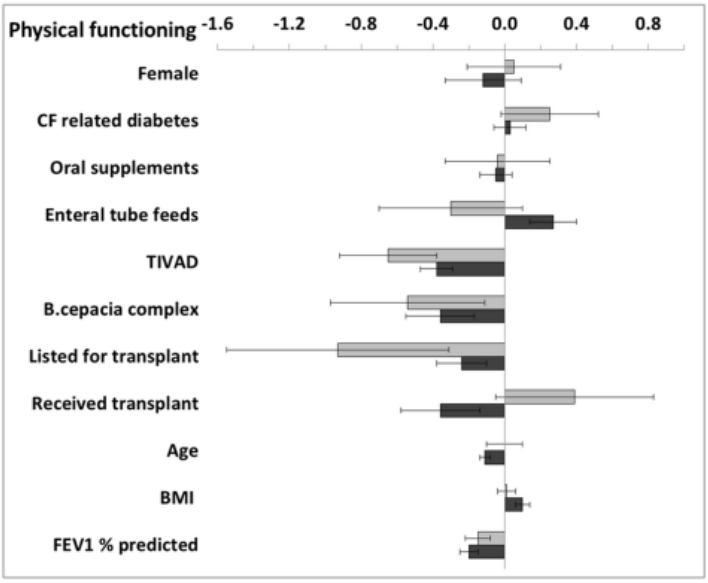

Figure 1. Model coefficients for the cross-sectional data (light grey) and the longitudinal data (dark grey) for the domain Physical functioning. Bars show plus and minus one standard error.

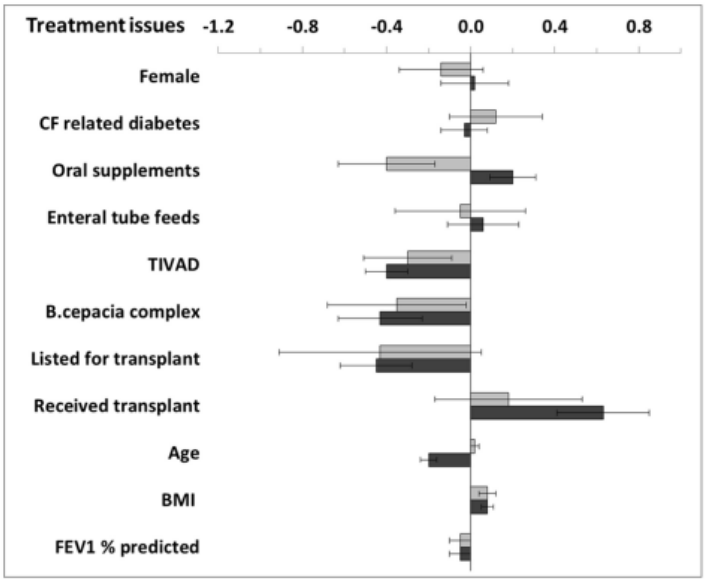

Figure 2. Model coefficients for the cross-sectional data (light grey) and the longitudinal data (dark grey) for the domain Treatment issues. Bars show plus and minus one standard error.

Not surprisingly, for both Physical functioning and Treatment issues, the standard errors were larger for the cross-sectional data than for the longitudinal data since the former were based on 182 assessments whereas the later were based on 642 assessments, albeit that the 642 assessments were not independent. For Physical functioning, the coefficients for 'received transplant' and for 'enteral tube feeds' changed sign and for Treatment issues, the coefficient for 'oral supplements' also changed sign. The effect of age was negligible for both Physical functioning and Treatment issues for the cross-sectional data but an age effect was clearly evident for both in the longitudinal data. The other coefficients were reasonably consistent between the two data sets within the bounds of uncertainty as expressed by the standard errors but the failure to detect and so adjust for the effect of increasing age was a most concerning failure of the cross-sectional data.

4. SIMULATION STUDY

When comparing the coefficients between the fitted models for the cross-sectional and the longitudinal data, the true coefficients were unknown quantities and hence a simulation study using known parameter values was performed to provide additional insight into the comparison. Firstly, M patients were simulated by selecting a bootstrap sample with replacement from the 182 patients. M took the values 200, 650, 1300, 2600, 5200, 10400 in turn. Secondly, all of the demographic and clinical variables at each assessment in sequence for a selected patient were entered into the simulated longitudinal data and flagged as belonging to the *i*th patient in the bootstrap sample. Thirdly, a HRQoL score, Q_{ij}, was calculated for the simulated patient i at time j. Q_{ij} was obtained by calculating p_{ij} using the linear equation:

$$\text{logit}(p_{ij}) = \beta_{0i} - \beta_1 G_i - (\beta_2/5)(\text{age}_{ij}\text{-}30) - \beta_3 C_{ij} - (\beta_{4i}/10)(\text{FEV}_1\%\text{predicted}_{ij} - 30)$$

where $\beta_{0i} \sim N(1.31,(0.2x1.31)^2)$, $\beta_{4i} \sim N(0.2,(0.2x0.2)^2)$ and Covariance(β_{0i}, β_{4i}) = 0, $\beta_1 = 0.12$, $G_i = 0$ when patient i was male and $G_i = 1$ when patient i was female, $\beta_2 = 0.11$ and age was measured in years, centred on 30 years and $\beta_3 = 0.36$, and $C_{ij} = 0$ when patient i did not have *B. cepacia* complex infection at time j and $C_{ij} = 1$ when patient i had *B. cepacia* complex infection at time j.

Explaining this procedure in more detail, a bootstrap sample of size 182 patients was selected randomly with replacement from the 182 actual patients. Let us say that person A was in this bootstrap sample and that person A had measurements at four times, these times denoted by T1, T2, T3 and T4. For T1 person A's age, gender, FEV_1% predicted and B. *cepacia* infection (present or absent) at time T1 were used to generate by simulation person A's QoL at time T1. Then the same

procedure was used for times T2, T3 and T4. If another person, person B was also in the bootstrap sample and person B had been measured at only two times, S1 and S2 then their recorded values of age, gender, FEV_1% predicted and B. *cepacia* infection at times S1 and S2 were used to simulate their QoL at times S1 and S2 respectively. Once this procedure was carried out for all 182 in the bootstrap sample a simulated data set was complete.

The standard deviations of the two normal distributions were chosen to provide a 20% coefficient of variation for the patient-specific random coefficients and the parameter scaling (5 years for age and 10 years for FEV_1% predicted) was chosen so that coefficients could be expressed to two decimal places in tables and figures. Age was centred on 30 years and FEV_1% predicted was centred on 30% because these values limited the correlation between the parameter estimates for the random coefficients in the models for the empirical data. Finally, once p_{ij} had been calculated, score Q_{ij} was obtained by generating a random value from a binomial distribution with probability p_{ij} and binomial denominator of either 50 or 15. The computation of Q_{ij} for each denominator used the same bootstrap samples of patients, but used independently generated values of the random coefficients β_{0i} and β_{4i} and independently generated values from the binomial distributions. Upon completion of a simulated longitudinal data set, a cross-sectional data set was extracted by using the first time point for each simulated patient. This method should generate a similar degree of correlation between estimated coefficients from the simulated cross-sectional and longitudinal data as was observed in the empirical data. The simulation was carried out in the R programming language (R Development Core Team, 2010).

A simplified HRQoL model was chosen for the simulation study to make results easier to interpret, but one of each type of variable was included. The constant term was a necessity and could not be excluded. FEV_1% predicted was known to be related to eight out of the nine domains (Abbott et al., 2013) and was a random patient-specific coefficient. Gender was a classifying factor which did not change for patients over time and the cross-sectional data might be expected to estimate a gender difference reasonably well. On the other hand, *B. cepacia* complex infection was present for some patients at the first point but other patients developed the infection at a later time point. If treated early and aggressively *B. cepacia* complex infection can be eradicated and so, if a patient has the infection at time *j* then they do not necessarily have the infection at subsequent times. The effect of *B. cepacia* complex infection on HRQoL in the longitudinal data would be established partly from the cross-sectional data subset and partly from the subsequent longitudinal data. Age was a variable which the empirical data suggested might be poorly estimated by cross-sectional data but which, a priori, was highly likely to be an important predictor of HRQoL and hence age was included in the

simulated models.

Table 5: Summary table for the simulation study showing (i) the longitudinal and cross-sectional numbers of patients needed to achieve a target standard error of 30% of the true coefficient value and (ii) the ratio of the numbers of patients required for a cross-sectional study compared to a longitudinal study to achieve the same coefficient standard errors.

Coefficient	True value	Target SE	Longitudinal sample size	Cross-sectional sample size	Ratio	95% CI	
Maximum=50							
Intercept	1.31	0.393	3	8	2.61	2.32	2.92
Gender	0.12	0.036	448	882	1.97	1.82	2.13
Age	0.11	0.033	35	129	3.63	3.30	4.02
B. cepacia comp.	0.36	0.108	114	260	2.28	1.93	2.68
FEV_1% predicted	0.20	0.060	6	16	2.70	2.31	3.16
Maximum=15							
Intercept	1.31	0.393	6	21	3.56	3.22	3.93
Gender	0.12	0.036	804	2363	2.94	2.73	3.15
Age	0.11	0.033	76	335	4.39	4.02	4.79
B. cepacia comp.	0.36	0.108	227	675	2.98	2.62	3.39
FEV_1% predicted	0.20	0.060	13	42	3.18	2.78	3.63
Ratio sample size							
Maximum=15 to Maximum=50			1.8 to 2.2	2.6 to 2.7			

Figure 3 shows the estimated coefficients from the simulated cross-sectional and longitudinal data for the increasing number of patients from 200 to 10,400. Any bias would seem to be equal in extent for the cross-sectional and longitudinal data. Figure 4 shows the estimated standard errors of the coefficients. Ordinary linear regression of the logarithm of the estimated standard errors on the logarithm of the simulated number of patients for each of the five variables and for both cross-sectional and longitudinal data strongly supported the hypothesis that the standard error of a coefficient was proportional to the square root of the number of patients, at least to a good approximation. Consequently, the standard errors for each coefficient were modelled by a simple linear regression which regressed y_k where $y_k = \log(SE_k) + 0.5\log(N_k)$ on the indicator variable D_k which took the value 0 if the standard errors were from a longitudinal study and 1 otherwise. This allowed the estimation of the multiplying factors by which we would need to increase the number of patients to obtain the same precision of estimation in a cross-sectional study as had been obtained from a longitudinal study. This, of course, assumed that HRQoL was a function of age, gender, *B. cepacia* complex infection and FEV1% predicted only. The ratios are shown in Table 5.

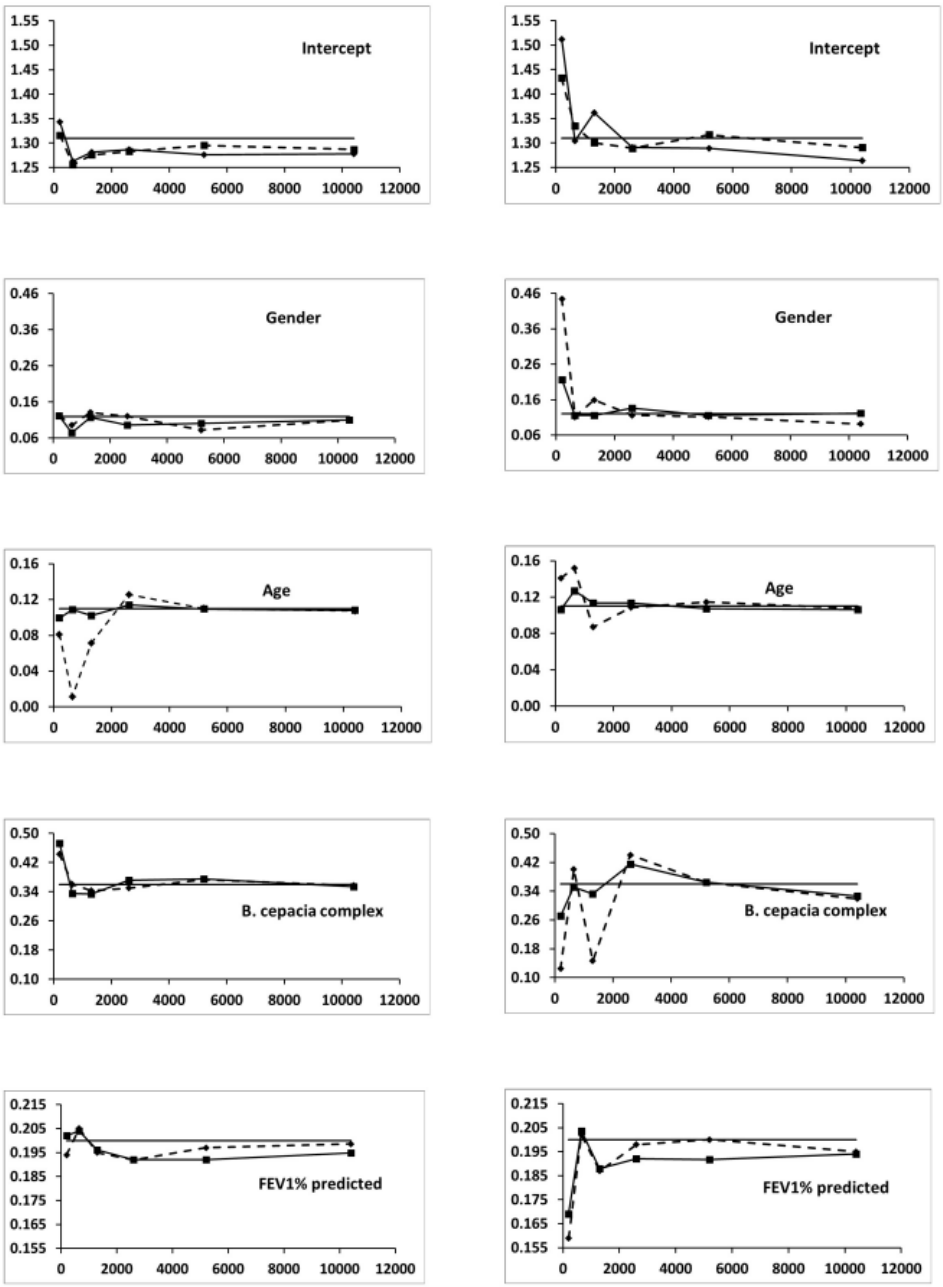

Figure 3. Estimated coefficients for the longitudinal simulated data (solid line with markers), for the cross-sectional simulated data (broken line with markers) and the true value (solid line without markers) against the simulated number of patients. Left-hand column is for a binomial denominator of 50 and-right hand column is for 15.

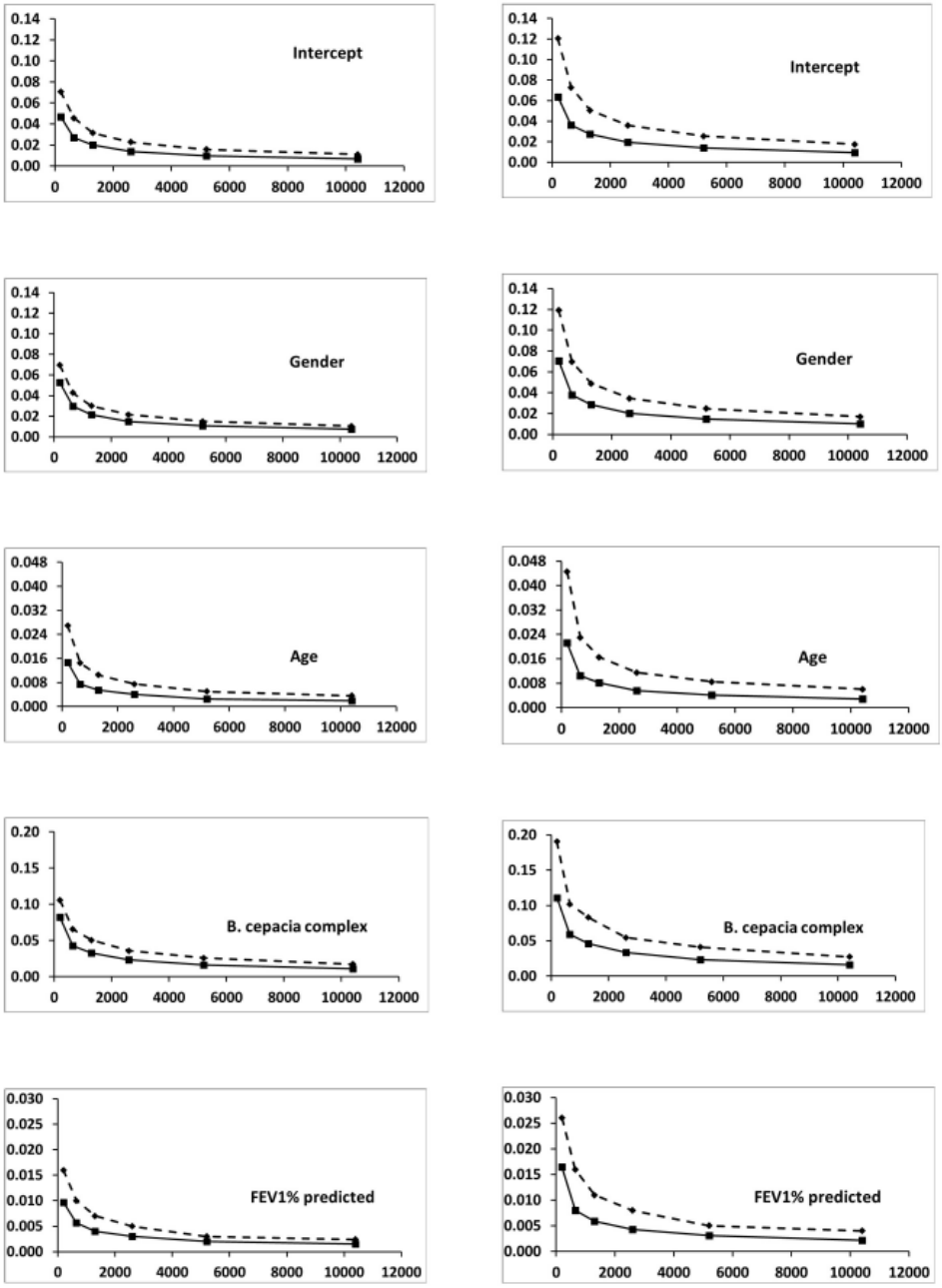

Figure 4. Coefficient standard errors for the longitudinal simulated data (solid line with markers), the cross-sectional simulated data (broken line with markers) against the simulated number of patients. Left-hand column is for a binomial denominator of 50 and right-hand column is for 15.

For example, if a longitudinal study had 100 patients, a HRQoL domain which was measured out of maximum score of 50 (and had other characteristics similar to this study) then a cross-sectional study would need 228 patients (100 x 2.28) to achieve the same precision in estimation of the effect of *B. cepacia* complex infection. For the age coefficient, 363 patients would be required in a cross-sectional study compared to 100 in a longitudinal study. If the HRQoL domain had a maximum score of 15 then the patients required for the cross-sectional study would be 298 and 439 respectively for every 100 patients in the longitudinal study. Table 5 also shows the number of patients needed in a longitudinal and a cross-sectional study to achieve a standard error of 30% of the true value. A standard error of about 30% should be adequate to detect significance of a clinical or demographic variable. Table 5 is therefore a summary of the information contained in Figures 3 and 4.

The results presented in Table 5 support the following general insights:

(i) A longitudinal study of a fixed number of patients will offer far less precision of estimation and lower power to detect the effect of clinical and demographic variables for a HRQoL domain with a small maximum (e.g.15) compared to a domain with a large maximum (e.g. 50). The same is true for a cross-sectional study. However, to obtain the same precision for a domain maximum of 15 as for a domain maximum of 50 the increase in numbers of patients needs to be by a factor of 1.8 to 2.2 for a longitudinal study. For a cross-sectional study the factor needs to be 2.6 to 2.7. So, it would seem likely that longitudinal studies better serve HRQoL domains with small maxima than cross-sectional studies.

(ii) Even in a longitudinal study, a relatively large number of patients might be required to obtain a precise estimate of the effect of a demographic variable such as gender which does not change longitudinally.

(iii) A longitudinal study with modest numbers of patients can estimate, with good precision, the effect of increasing age. To estimate age with the same precision in a cross-sectional study the number of patients would need to be multiplied by 3.63 for domain maximum of 50 and 4.39 for domain maximum of 15. Therefore, a cross-sectional study requires relatively more assessments than a longitudinal study to achieve the same precision for the effect of age.

(iv) The estimation of the effect of a variable such as *B.cepacia* complex infection which does change with time would seem to be equally well estimated by a cross-sectional survey as by a longitudinal survey with similar numbers of assessments.

5. DISCUSSION

The empirical demographic and clinical data used for the simulations had 642 patient-assessments for 182 patients providing an average of 3.5 assessments per patient. However, the 642 assessments from 182 patients in a longitudinal study may contain less or possibly even more information about a parameter than 642 assessments from 642 patients in a cross-sectional study. The longitudinal study may contain less information because of the correlation between observations on the same individual for whom assessments would not be independent. On the other hand, the longitudinal study monitoring the same individual over a range of ages may be better able to detect the response in HRQoL to change in demographic (age) and clinical variables at the individual level. The ratios given in Table 5 are a measure of whether the longitudinal or cross-sectional study performs better for each variable since these can be compared with the bench mark value of 3.5. A higher ratio goes in favour of a longitudinal study and a lower ratio favours the cross-sectional study. Looking at how the ratios change as the maximum score reduces would indicate that a cross-sectional study may be suitable when the maximum is high but a longitudinal study rapidly gains advantage for HRQoL domains with small maxima. Also the large ratio required for age may explain why the empirical longitudinal and cross-sectional studies gave such discrepant results for age.

The sample sizes required to estimate parameter values to within a 30% precision were surprising. A longitudinal sample size of 182 with, on average, 3.5 assessments per patient should be adequate to detect the significant effect of age, *B. cepacia* complex and FEV_1% predicted, if these were the only significant variables, for domains with maxima in the range 15 to 50. However, the effect of gender requires a considerably larger number of patients in a longitudinal study for an effect to be detected. The effect of gender seems partially lost and this might be due to the inclusion in the model of a random intercept for each patient. Intuitively a random intercept should be included in the models and tests of whether the intercept should be a random or fixed effect for the empirical data indicated that a random intercept was required. These were tests for whether the intercept variance and the covariance with other random effects were not significantly different to zero. With a random intercept included, the failure to detect the effect of gender ought not to bias judgements regarding the effect of the other variables.

Sample sizes to detect significant effects would clearly need to be larger if more than four variables were included in the simulations and in the fitted models because standard errors would be increased. If more variables were included in the model simulations but were not included in the fitted model this would appear as

extra binomial variation and it is not clear what influence this would have on ratios and sample sizes. Most likely sample sizes would increase but the influence on the ratios is not predictable. Also, the simulations of HRQoL contained two random coefficients; the intercept and the gradient with FEV_1% predicted. The two random coefficients were generated uncorrelated in the simulations and the value, on which FEV_1% predicted was centred, was selected to reduce correlation in the parameter estimates. If some modest correlation was simulated then, again, the effect on ratios and samples sizes would be difficult to predict.

The way in which the HRQoL was simulated may have impacted on the sample sizes and ratios with which the two types of study have been compared. As an alternative, the logit of the HRQoL binomial probability could have been generated using the method described here but only for the first time point for each patient. For later time points, the binomial probability value at the previously recorded time point could have been adjusted by the change in demographic and clinical variables since the earlier time point. This would have introduced additional autocorrelation into the longitudinal data. The method used in the simulations introduced autocorrelation in the HRQoL only through the autocorrelation in the demographic and clinical variables. Again it is not possible to predict what difference this would have made to the sample sizes and ratios. More work could provide some answers.

All of the foregoing discussion has been concerned with statistical issues but clearly the choice between a longitudinal study and a cross-sectional study is not determined by statistics. There are many factors which influence the choice, not least of which is the time scale under which funding is provided for research. Few funders would fund a 12 year prospective study and the empirical study described here was unfunded research. By necessity the study was single centred and had less than 250 patients in total and was manageable within the staff time and resources available. In contrast, the International Depression/anxiety Epidemiological Study (TIDES) in the UK was a cross-sectional study which recruited about 45% of the UK adult population in the CF Registry across 25 adult CF centres (Duff et al., 2014). This large study recruited a total of 1780 adults and was a major undertaking over a 30 month period. It would have been unlikely to have recruited more had it continued beyond 30 months because in this type of study a saturation point can be reached with a finite population of potential patients. A cross-sectional study, therefore, of a particular patient group will be limited above by the population size of the group and the willingness of the patients to participate in research. Statisticians can simulate 10,000 patients but other researchers can only dream of obtaining such sample sizes. The number of patients required for a cross-sectional study equivalent

to a given longitudinal study might simply not be achievable. A longitudinal study of a small number of patients will be bound to yield more information than a cross-sectional study of an equally small number of patients. In this case, the researchers should choose a longitudinal study and cultivate a lot of 'patience' rather than a lot of 'patients'.

REFERENCES

Abbott, J. (2009). Health-related quality of life measurement in cystic fibrosis: advances and limitations. In *Chronic Respiratory Disease* 6: 31-41.

Abbott, J., Hart, A., Havermans, T., Matossian, A., Goldbeck, L., Barreto, C., Bergsten-Brucefors, A., Besier, T., Catastini, P., Lupi, F. and Staab, D. (2011). Measuring health-related quality of life in clinical trials in cystic fibrosis. In *Journal of Cystic Fibrosis* 10: S82-S85.

Abbott, J., Hurley, M.A., Morton, A.M. and Conway, S.P. (2013). Longitudinal association between lung function and health-related quality of life in cystic fibrosis. In *Thorax* 68: 149–154.

Abbott, J., Morton, A.M., Hurley M.A. and Conway, S.P. (2015) Longitudinal impact of demographic and clinical variables on health-related quality of life in cystic fibrosis. In *BMJ Open* 5(5), doi:10.1136/bmjopen-2014-007418.

Duff, J.A., Abbott, J., Cowperthwaite, C., Sumner, C., Hurley, M.A., Quittner, A. and TIDES-UK Group. (2014). Depression and anxiety in adolescents and adults with cystic fibrosis in the UK: A cross-sectional study. In *Journal of Cystic Fibrosis* 13: 745-753.

Gee, L., Abbott, J., Conway, S.P., Etherington, C. and Webb, A.K. (2000). Development of a disease specific health related quality of life measure for adults and adolescents with cystic fibrosis. In *Thorax* 55: 946-54.

Gee, L., Abbott, J., Conway S.P., Etherington, C. and Webb, A.K. (2003). Quality of life in cystic fibrosis: The impact of gender, general health perceptions and disease severity. In *Journal of Cystic Fibrosis* 2: 206-213.

Gee, L., Abbott, J., Hart, A., Conway S.P., Etherington, C. and Webb, A.K. (2005). Associations between clinical variables and quality of life in adults with cystic fibrosis. In *Journal of Cystic Fibrosis* 4: 59-66.

Goldbeck, L., Zerrer, S. and Schmitz, T.G. (2007). Monitoring quality of life in outpatients with cystic fibrosis: Feasibility and longitudinal results. In *Journal of Cystic Fibrosis* 6: 171-178.

Goldstein H. (2011). *Multilevel Statistical Models,* 4th Edition. Wiley, Chichester.

Laborde-Castérot , H., Donnay, C., Chapron, J., Burgel, P.R., Kanaan, R., Honoré, I., Dusser, D., Choudat, D. and Hubert, D. (2012). Employment and work disability in adults with cystic fibrosis. In *Journal of Cystic Fibrosis* 11:137-43.

Rasbash, J., Steele, F., Browne, W.J. and Goldstein, H. (2009). *MLwiN Version 2.10*, Centre for Multilevel Modelling, University of Bristol, Bristol.

R Development Core Team (2010). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org. Last access 11th January 2016.

Rindfleisch, A., Malter, A.J., Ganesan, S. and Moorman, C. (2008). Cross-sectional versus longitudinal survey research: concepts, findings and guidelines. In *Journal of Marketing Research* 45: 261-279.

Rosenstein, B.L. and Zeitlin, P.L. (1998). Cystic fibrosis. In *Lancet* 351: 277-282.

Sawicki, G.S., Rasouliyan, L., McMullen, A.H., Wagener, J.S., McColley, S.A., Pasta, D.J. and Quittner, A.L. (2011). Longitudinal assessment of health-related quality of life in an observational cohort of patients with cystic fibrosis. In *Pediatric Pulmonology* 46: 36-44.

UK Cystic Fibrosis Registry (2014). *Annual Data Report 2013*. Cystic Fibrosis Trust.

Wahl, A.K., Rustoen, T., Hanestad, B.R., Gjengedal, E. and Moum, T. (2005). Living with cystic fibrosis: Impact on global quality of life. In *Heart and Lung* 34: 324-331.

Walters, S.J. (2009). *Quality of Life Outcomes in Clinical Trials and Health Care Evaluation: A Practical Guide to Health Care Interpretation*. Wiley, Chichester.