Understanding the Natural Progression in %FEV$_1$ Decline in Patients with Cystic Fibrosis

D Taylor-Robinson[1], M Whitehead[1], HV Olesen[4], T Pressler[5], RL Smyth[2] , P Diggle[3]

Division of Public Health and Policy[1], School of Reproductive and Developmental Medicine[2], University of Liverpool, CHICAS - Lancaster University[3] UK, , Aarhus University Hospital, Aarhus[4], Rigshospitalet, Copenhagen, Denmark[5]

Background

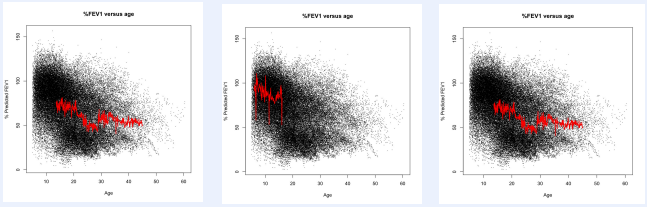
%FEV$_1$ is a key outcome in cystic fibrosis, and is one of the most important predictors of death.

Currently used modelling approaches have limitations when applied to datasets with repeated %FEV$_1$ measures over longer follow-up periods.

We describe a novel modelling approach for analysing changes in %FEV$_1$ over time, and apply this to a unique population level dataset from Denmark to describe the correlation in %FEV$_1$ in individuals over time, and the effect of covariates.

Unique Danish dataset with frequent measures

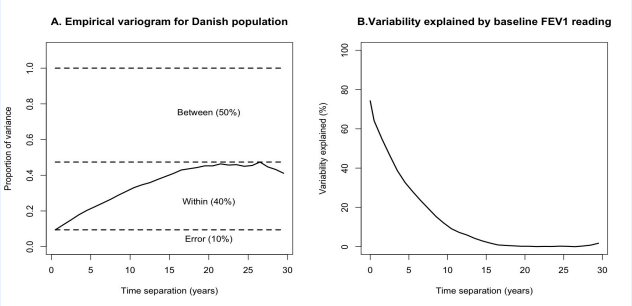
The analysis uses longitudinal %FEV1 data from the Danish CF population collected over long follow up periods. The dataset contains 70,448 measures on 479 patients seen between 1969 and 2010, with a total of 6500 person-years of follow up. The long follow-up presents modelling challenges.



Plots of %FEV1 versus age with 3 individual trajectories highlighted

Quantifying the variability in %FEV$_1$ with the variogram approach

Panel A shows the scaled empirical variogram for the Danish data. The solid line (variogram function) represents the variance of the difference between residual errors within individuals at time lags from zero to 30 years. The variogram function increases up to about 20 years, corresponding to a decreasing correlation between paired lung function measures with increasing time-separation. The variogram partitions the variability in the data into three components: within person, between person, and error. Panel B shows the proportion of variability in an individual's %FEV$_1$ that is explained by their %FEV$_1$ at baseline.



Modelling framework

We use a linear mixed effects model with longitudinally structured correlation. Model fitting involves:

- Fitting a provisional model for the mean response by ordinary least squares (OLS)
- Using the empirical variogram of the residuals (Diggle 2002) to provide initial estimates for the three components of variation (error, within person, between person), and for the shape of the correlation function of the between-times-within-subjects component
- Re-estimation of all of the model parameters by maximum likelihood estimation. We can then assess associations between single or multiple covariates and the population mean %FEV1 over time.

Let Y_{ij} denote the jth repeated %FEV1 measurement on the ith patient:

$$Y_{ij} = \mu_{ij} + U_i + W_i(t_{ij}) + Z_{ij}$$

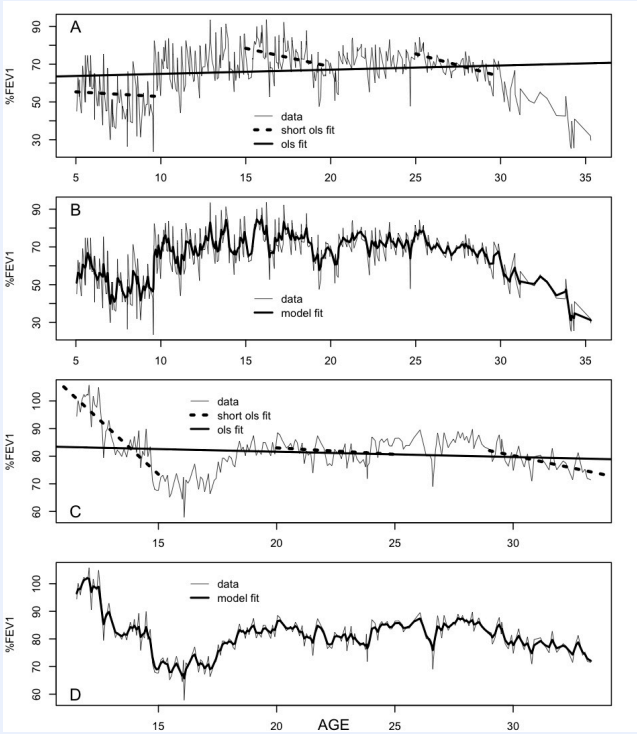
μ_{ij} is the mean, population-averaged, response specified as a linear combination of explanatory variables

U_i describes how the average lung function of the ith patient varies about the population-averaged response for all patients with the same values of the explanatory variables x_{ijk}

$W_i(t)$ are independent copies of a stationary Gaussian process with mean zero, variance σ^2 and correlation function $\rho(u) = Corr\{Y_i(t), Y_i(t-u)\}$.

Exponential correlation function, $\rho(u) = \exp(-|u|/\phi)$, in which the parameter ϕ describes the rate at which the correlation decays towards zero with increasing time-separation, u. Z_{ij} is measurement error.

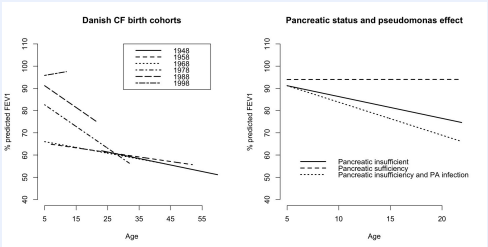
Stationary Gaussian process model fits better than random intercept and slope



Panel A shows the data for a single individual, illustrating that a linear trend fits reasonably well over short time-periods, but gives a very poor fit over long periods; Panel B shows the same data with the fitted trajectory of the stationary Gaussian process model; the fitted trajectory is a smoothed version of the observed %FEV1 trajectory; the amount of smoothing is determined as an automatic by-product of the fitting process so as to deliver uncorrelated residual variation. Panels C and D show the corresponding plots for a second individual.

The model fit line better represents underlying lung function and could inform clinical decision making

Using the model to explore the effect of cohort, pseudomonas and pancreatic status



Pseudomonas infection and pancreatic insufficiency are associated with significantly increased rates of lung function decline

Conclusions

-Significant short term variability in %FEV$_1$ measures

-Long term correlation in %FEV1 measures, such that an individuals measure in early life is correlated up to 15 years subsequently

significant cohort effect, with improvement in contemporary cohorts

-Pseudomonas acquisition and pancreatic insufficiency associated with increased rate of decline of %FEV1