UNIVERSITY OF GOTHENBURG

MASTER'S THESIS

Joint mixed-effects modelling of longitudinal health data for monitoring the progression of cystic fibrosis lung disease

Mattis Hallberg

UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2024-2025
www.gu.se

Thesis for the Degree of Master of Science

Joint mixed-effects modelling of longitudinal health data for monitoring the progression of cystic fibrosis lung disease

Mattis Hallberg

UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Division of Mathematical Statistics
SE – 412 96 Gothenburg, Sweden
Gothenburg, May 2025

Joint mixed-effects modelling of longitudinal health data for monitoring the progression of cystic fibrosis lung disease
Mattis Hallberg

Supervisor: Henrik Imberg, Statistiska Konsultgruppen Sweden and Department of Molecular and Clinical Medicine, University of Gothenburg.
Examiner: Johan Jonasson, Division of analysis and probability theory, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg.

Master's Thesis 2025
Department of Mathematical Sciences
University of Gothenburg

Typeset in LaTeX

Abstract

Cystic fibrosis (CF) is a genetic disease that affects several organs of the body such as the liver, pancreas, or the lungs. Cystic fibrosis is monitored using hospital and home spirometry, where the difference between the two is what type of technique that is used. This study evaluates joint mixed-effects models as a statistical framework for analysing home and hospital spirometry measurements. Statistical methods, such as simple linear regression and univariate mixed-effects models can be used to analyse spirometry. However, both simple linear regression and univariate linear mixed-effects models considers only hospital spirometry and do not accommodate the rich patient-generated data obtained by home spirometry. By jointly modeling home and hospital spirometry, we aim to investigate prediction accuracy of hospital spirometry. We compare the performance of joint and univariate linear mixed-effects models with that of simple linear regression, evaluating their ability to predict: (i) the current (i.e., most recent) hospital spirometry value for a patient, (ii) the current mean value, (iii) the subject-specific trend, and (iv) the population-average trend. The results showed that joint linear mixed-effects models and univariate linear mixed-effects models should be used over simple linear regression to estimate trends and specific values in a CF-setting. Moreover, the results supported the use of joint linear mixed-effects models when jointly modeling home and hospital spirometry data.

Keywords: FEV_1; $FEV_1\%$; Home monitoring; Linear mixed-effects models; Joint model; Longitudinal data; Lung function; Remote healthcare; Spirometry.

Acknowledgements

I would like to thank my supervisor, Henrik Imberg at Statistiska Konsultgruppen, for guidance and helping me move in the best direction during this project. I would like to thank Marcus Svedberg for providing literature and showing me the CF clinic in Gothenburg and thanks to Vilhelm Larsson, Statistiska Konsultgruppen, for contributing with help and ideas. Finally, I want to thank my parents for always supporting and believing in me.

Mattis Hallberg, Gothenburg, May 2025

List of abbreviations

BLUE	Best linear unbiased estimator
BLUP	Best linear unbiased predictor
CF	Cystic fibrosis
CI	Confidence interval
CM	Current mean
CV	Current value
FEV_1	Forced expiratory volume in one second
$FEV_1\%$	Percent predicted forced expiratory volume in the first second
FVC	Forced vital capacity
GLI	Global lung function initiative reference equations
JLMM	Joint linear mixed-effects model
LMM	Linear mixed-effects model
LOOCV	Leave-one-out cross-validation
MME	Mixed-model equations
OLS	Ordinary least squares
PAT	Population-average trend
RMSE	Root mean square error
RMSE(CM)	RMSE of current mean predictions.
RMSE(CV)	RMSE of current value predictions.
RMSE(PAT)	RMSE of population-average trend predictions.
RMSE(SST)	RMSE of subject-specific trend predictions.
SLR	Simple linear regression
SST	Subject-specific trend
ULMM	Univariate linear mixed-effects model

Nomenclature

Indices

i	Index for participant (individual).
j	Index for measurement occasion within a participant.
k	Index for measurement setting, where $k \in \{1, 2\}$; $k = 1$ denotes home spirometry and $k = 2$ denotes hospital spirometry.

Parameters

n	Total number of participants in the sample or simulation.
t_{ij}	Time of measurement j for participant i; typically represented by age or visit time.
$\beta_0^{(k)}$, $\beta_1^{(k)}$	Fixed effects: global intercept and slope for setting k.
$b_{0i}^{(k)}$, $b_{1i}^{(k)}$	Random effects: participant-specific intercept and slope for setting k.
$\beta_0^{(k)} + b_{0i}^{(k)}$	True intercept for participant i in setting k.
$\beta_1^{(k)} + b_{1i}^{(k)}$	True slope for participant i in setting k.
$\hat{\beta}_0^{(k)} + \hat{b}_{0i}^{(k)}$	Estimated intercept for participant i in setting k.
$\hat{\beta}_1^{(k)} + \hat{b}_{1i}^{(k)}$	Estimated slope for participant i in setting k.
$\hat{b}_{0i}^{(k)}$, $\hat{b}_{1i}^{(k)}$	Predicted random-effects intercept and slope parameter for participant i in setting k from the full data.
$\hat{b}_{0i,-j}^{(k)}$, $\hat{b}_{1i,-j}^{(k)}$	Predicted random-effects intercept and slope parameter for participant i in setting k, fitted after removing measurement j.
$\tilde{\beta}_{0,-j}$, $\tilde{\beta}_{1,-j}$	Estimated fixed-effects intercept and slope parameter after excluding measurement j from participant i.

Variables

$Y_{ij}^{(k)}$	Random variable representing the spirometry value (e.g., FEV1%) for participant i at measurement j in setting k.
Age_{ij}	Age of participant i at measurement j.
ε_{ij}	Random error term (residual) for participant i at measurement j.
$y_{ij}^{(k)}$	Observed spirometry value for participant i at measurement j in setting k.
$\hat{y}_{ij}^{(k)}$	Predicted spirometry value for participant i at measurement j in setting k, estimated from the full data.
$\hat{y}_{i,-j}^{(k)}$	Predicted spirometry value for participant i at measurement j in setting k, estimated after removing measurement j.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Cystic fibrosis (CF) is a genetic disease that can affect several organs of the body such as the liver, pancreas, or the lungs. CF is often characterized by progressive lung function decline [28], and therefore there is a need to regularly monitor the progression of the disease. Some techniques that are used to monitor CF lung disease are chest X-rays, computed tomography, and hospital-based spirometry ([27], [26], [30]). The technique that is most commonly used is spirometry and it is used to assess lung function by measuring how much and how quickly air can be exhaled.

People with CF often face the challenge of traveling considerable distances to access the specialised CF clinics to receive healthcare [20]. The appointments can disrupt various aspects of the patient's life, such as missing school or work, and other extracurricular activities. To reduce the burden of in-person visits, remote monitoring using home spirometry has become increasingly common in CF care.

Home spirometry is performed by using a portable device, usually with turbines or ultrasonic technique's, and hospital spirometry uses the pneumotachometer. Combining home and hospital spirometry has been proven to be effective in the monitoring of the disease and it improves the remote healthcare [20]. Also, weekly follow-ups of home spirometry is a reliable tool for telemonitoring of CF lung disease [4]. Interpretation of spirometry values should be complemented with clinical expertise to make a correct assessment of the individual's lung function measurements [25]. Current guidelines recommend hospital spirometry every third month, but the optimal frequency remains unclear. Home spirometry is flexible and it is possible to share the data with healthcare clinics but earlier studies ([22], [5], [10]) have raised concerns about the reliability of home spirometry due to conflicting results regarding the interchangeability of hospital and home spirometry measurements. Being able to replace hospital measurements with home measurements could improve the remote healthcare further but the reliability of the monitoring techniques cannot be compromised.

The longitudinal nature of spirometry data necessitates statistical methods that account for within-subject correlation over time. Longitudinal data consists of repeated observations of the same variable over time. This is where linear mixed-effects models (LMMs) are particularly useful, as they are a natural and well-suited choice for analyzing longitudinal data. LMMs can assess trends on both a population level and on an individual level and they are often used to model only one response variable over time. However, by only modeling one response variable and not several correlated response variables simultaneously over time could potentially mean that valuable information is lost. In this case, the correlation between the variables could potentially be used by the model to improve predictions.

Joint mixed-effects models extend traditional LMMs by simultaneously modeling multiple correlated outcomes, allowing for more efficient use of available data. By combining self-generated data (e.g., home spirometry) and healthcare-generated data (e.g., hospital spirometry) into the same analytical model, the correlation between the two outcomes could improve accuracy in estimation and increase the understanding of the progression of CF lung disease. For example, Ivanova [15] uses linear mixed-effects models to model several correlated outcomes in a non CF context. The purpose of this study is to utilize joint linear mixed-effects models to model home and hospital spirometry as two correlated outcomes in order to make predictions of hospital spirometry. The model is tested through simulations in different theoretical scenarios to investigate how the prediction accuracy is affected and compared to univariate linear mixed-effects models and simple linear regression.

1.2 Aim

The aim of this thesis is to evaluate the utility of joint mixed-effects models for analyzing longitudinal data with two correlated outcomes, such as measurements collected in different settings, with applications to lung function monitoring in cystic fibrosis (CF). Specifically, the thesis investigates whether combining data sources within a joint modeling framework can improve prediction accuracy compared to univariate mixed-effects models and simple linear regression.

1.3 Specific research questions

The specific research questions are as follows:

1. How does joint linear mixed-effects models compare to univariate mixed-effects models and simple linear regression in terms of accuracy across different time horizons?
2. How is model accuracy affected by the number of participants included in the analysis?
3. To what extent does the performance of the proposed methodology depend on the correlation between subject-specific trends?
4. Can measurements from one setting (e.g., hospital spirometry) be replaced by additional observations from another (e.g., home spirometry), without compromising predictive accuracy?

In addition, the thesis explores the following supplementary question:

- How stable are the different modeling approaches in terms of convergence and singularity when applied to realistic data scenarios?

1.4 Limitations

This project focuses exclusively on spirometry data, specifically, home- and hospital-based measurements, for individuals with CF. Other types of clinical or self-generated data, such as imaging-based indicators of structural lung damage (e.g., chest X-ray or computed tomography), growth metrics (e.g., weight, height, and BMI), antibiotic use, and patient-reported outcomes (e.g., CF symptoms and quality of life), are not considered in the current analysis.

The modeling approach is restricted to joint longitudinal data with continuous outcomes and relies on standard linear mixed-effects models. Assumptions include Gaussian-distributed random effects and homoscedastic residuals. Non-linear models, non-Gaussian distributions, and models with heteroscedastic variance structures fall outside the scope of this thesis.

Finally, the thesis focuses primarily on empirical performance through simulation experiments and applications, rather than the theoretical foundations of mixed-effects models. While a theoretical overview of mixed-effects modeling is provided in Chapter 2, no new methodological developments or in-depth theoretical derivations are presented. Instead, the emphasis lies on understanding the comparative performance and practical implications of joint versus univariate modeling frameworks.

1.5 Outline

The remainder of this thesis is structured as follows: Theory, Simulation experiments, Application results, and Discussion. The theory chapter provides and overview of the theory used, while simulation experiments discusses the implementation techniques and the results of these. Third, the application chapter discusses the implementation techniques and results applied on a real world CF data set. Lastly, the final chapter goes through the main findings, interpretations and explanations, clinical and practical implications, methodological considerations and future work, and conclusion.

2

Theory

2.1 Motivating example

Linear mixed-effects models are suited for longitudinal data that are hierarchical, correlated, and unbalanced, characteristics that are commonly found in longitudinal clinical studies of chronic diseases such as CF. In this thesis, the data consist of two types of spirometry measurements, i.e., home and hospital spirometry, repeatedly recorded by individuals with CF over time. Home and hospital spirometry measurements are inherently correlated, as both reflect lung function of the same individual. Furthermore, multiple measurements within each individual are correlated due to the longitudinal nature of the data. CF data have several common characteristics such as: the number of measurements varies substantially across individuals and some patients have frequent home and hospital recordings (especially home spirometry), while others have only a few. Measurements are taken at irregular time points which can be thought of as heterogeneous data with varying numbers of measurements collected at different time points. For people with CF, lung function typically declines over time, and this pattern is reflected in the data, where most individuals exhibit a downward trend in their spirometry values. In the modeling framework, each individual's lung function trend is assumed to be distributed around an overall population trend.

Figure 2.1 mimics a spirometry data set. The red lines represent the home spirometry lines for each participant and the blue are the hospital spirometry lines. The black line represents the population trend line of the home spirometry data and the dashed line is the population trend line of the hospital data. Note that, this is a toy example and it was used for conceptual demonstration.

The remainder of this chapter introduces the statistical foundations of LMMs and their extension to joint modeling frameworks, setting the stage for the simulation experiments and real-data applications that follow.

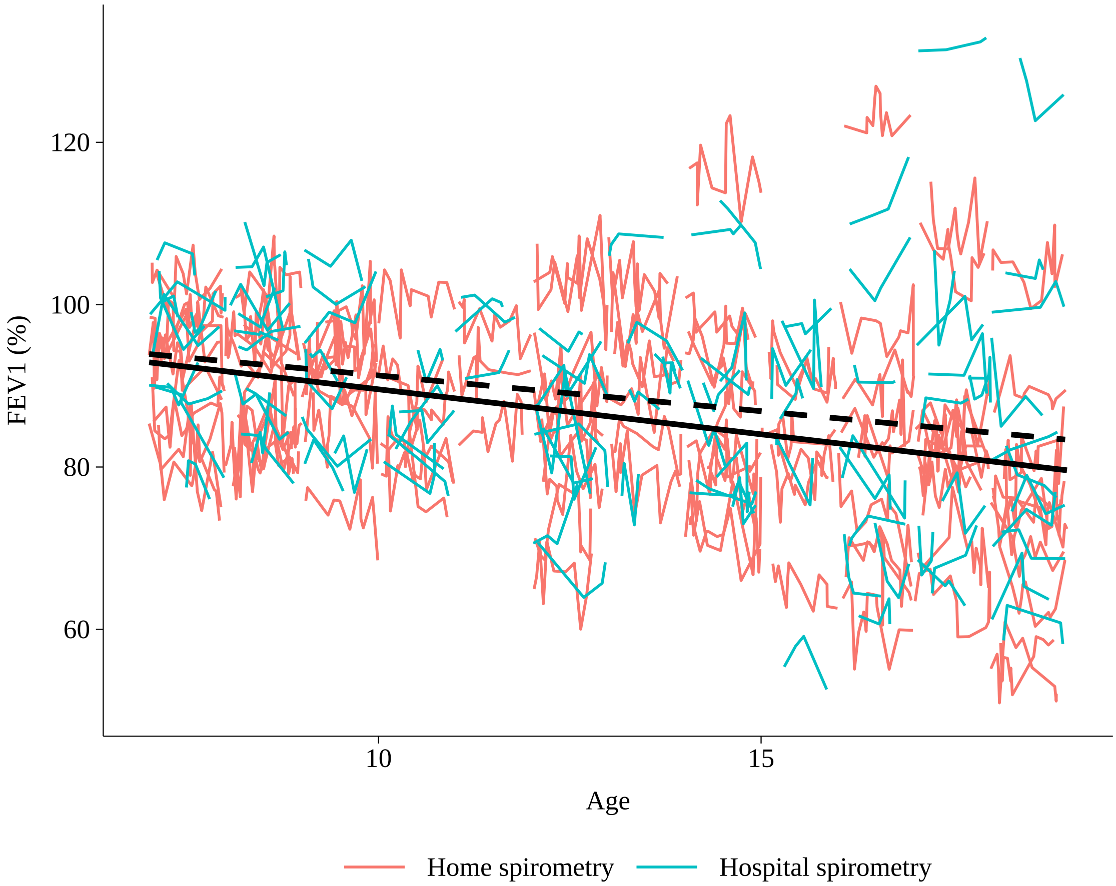

Figure 2.1: Illustrative example of simulated longitudinal spirometry data for multiple hypothetical individuals, showing home (red) and hospital (blue) measurements over time. The black line represent the population mean regression line of home spirometry and the dashed black line represent the mean regression line of hospital spirometry.

2.2 Linear mixed-effects models

To model CF data one could use linear mixed-effects models, which is an extension of simple linear regression (SLR). We begin by briefly reviewing the principles of simple linear regression. Simple linear regression is a method used to model the relationship between two variables: the predictor (explanatory variable) and the outcome (response variable). The model assumes a linear relationship, expressed as

$$Y = \beta_0 + \beta_1 x + \varepsilon, \tag{2.1}$$

where Y is the response, x is the predictor, β_0 is the intercept, β_1 is the slope, and ε is the error term. The goal is to estimate the parameters β_0 and β_1 such that the predicted values of Y are as close as possible to the observed values. Note, that the parameters that are to be estimated are called fixed parameters or fixed effects.

Fixed effects are effects that are assumed to be the same across all observations.

Linear mixed-effects models (LMMs) are an extension of SLR. LMMs are called mixed because they include both fixed effects and random effects. Random effects are allowed to vary between individuals and they follow some distribution, usually the normal distribution. Random effects usually occur when there are groups in the data. Linear mixed-effects models "borrows" information from the population trend to estimate individual trends, with shrinkage toward the population trend, especially for individuals with limited data. This yields more accurate estimates (best linear unbiased predictor) if the model assumptions are correct.

The random effects represent the deviation from the population trend for each individual. The random effects provides unique parameter estimates for each individual and the random effects takes into account the correlation between measurements within individual. Longitudinal data is different from independent measurements in the sense that repeated measures from the same individual are dependent. This makes a linear regression inappropriate due to the violation of independence.

The following section outlines the theoretical framework, progressing from the definition to parameter estimation. Section 2.2.1 discusses the general definition and how maximum likelihood techniques are used in the estimation procedure. Section 2.2.2 and 2.2.3 discusses how the fixed effects are estimated and how the random effects are predicted. Another technique to obtain parameter estimates are with the so called Henderson's mixed model equations, which are addressed in Section 2.2.4. The theory is based on Applied Mixed Models in Medicine by Brown and Prescott [13] and Linear Mixed Models: A Practical Guide Using Statistical Software by West, Welch, Galecki [32].

2.2.1 General definition

The general matrix definition of a linear mixed-effects model is:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Zb} + \boldsymbol{\varepsilon} \tag{2.2}$$

where $\mathbf{Y}$ is a random vector of responses, $\mathbf{X}$ is the design matrix of the fixed effects, $\mathbf{Z}$ is the design matrix of random effects, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed effects, $\mathbf{b}$ is the $q \times 1$ vector of random effects where $\mathbf{Cov}(\boldsymbol{b}) = \boldsymbol{G}$ and $\boldsymbol{b} \sim N(\mathbf{0}, \boldsymbol{G})$ accounting for subject specific variations, $\boldsymbol{\varepsilon}$ is the vector of residuals. Let $\mathbf{Cov}(\boldsymbol{\varepsilon}) = \boldsymbol{R}$. Note that p is the number of fixed effects and q is the number of random effects. Let $\boldsymbol{V}$ denote the covariance matrix of $\boldsymbol{Y}$ so that

$$\begin{aligned}
\boldsymbol{V} = \mathbf{Cov}(\boldsymbol{Y}) &= \mathbf{Cov}(\boldsymbol{X\beta} + \boldsymbol{Zb} + \boldsymbol{\varepsilon}) \\
&= \mathbf{Cov}(\boldsymbol{X\beta}) + \mathbf{Cov}(\boldsymbol{Zb}) + \mathbf{Cov}(\boldsymbol{\varepsilon}) \\
&= \boldsymbol{Z}\mathbf{Cov}(\boldsymbol{b})\boldsymbol{Z}^T + \mathbf{Cov}(\boldsymbol{\varepsilon}) \\
&= \boldsymbol{ZGZ}^T + \boldsymbol{R}
\end{aligned}$$

Here we use the fact that the covariance of $\boldsymbol{X\beta}$ is zero because of the deterministic effect and the random effects and residuals are uncorrelated. The general definition has the data of all individuals "stacked" upon each other. Observation between individuals are assumed to be independent and hence $\boldsymbol{V}$ is a block-diagonal matrix.

One could be more specific and specify it on individual level. The formula at individual level is

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i$$

and

$$\mathbf{Cov}(\boldsymbol{Y}_i) = \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i^T + \boldsymbol{R}_i.$$

This is not a diagonal matrix in general because measurements within a participant are correlated.

The estimation techniques that can be used are maximum likelihood (ML) and residual maximum likelihood (REML). Maximum likelihood produces downward biased estimates of the variance parameters [13] and REML is a modification of ML that handles this issue. This correction helps mitigate downward bias in variance estimates that arise from treating fixed effects as known when they are not.

Since the observations within an individual are dependent, a multivariate normal distribution for $\mathbf{Y}$ is used, where the random effects has zero mean and the distribution has mean vector $\boldsymbol{X\beta}$ and covariance matrix $\mathbf{V}$. Based on the normal assumption, the likelihood function is

$$\mathrm{L} = \frac{\exp\left\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{X\beta})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right\}}{(2\pi)^{n/2} \mid \boldsymbol{V} \mid^{1/2}}$$

and the corresponding log likelihood

$$\log(\mathrm{L}) = K - \frac{1}{2}\left[\log \mid \boldsymbol{V} \mid + (\boldsymbol{Y}-\boldsymbol{X\beta})^T\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{X\beta})\right], \tag{2.3}$$

where K is a constant.

The method of maximum likelihood is based on the concept of maximizing the log likelihood function with respect to the variance parameters while keeping the fixed effects constant. After the variance parameters are estimated, the fixed effects can be estimated by considering the variance parameters as constant and maximize the log likelihood with respect to the fixed effects, $\boldsymbol{\beta}$. Note that there are no closed formulas for the variance parameters and they are therefore estimated iteratively. A common algorithm used for this is the Newton-Raphson algorithm. Finally, the ML estimates are biased downwards and thus, the method of residual maximum likelihood can be used (REML).

First, let γ denote the variance parameters, i.e., the parameters in $\boldsymbol{V}$. It can be shown that $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$ are independent [13], hence the joint likelihood for $\boldsymbol{\beta}$ and γ can be expressed as a product of likelihoods. The likelihood for γ based on $\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$ is

$$\mathrm{L}(\boldsymbol{\gamma};\ \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) = \frac{\mathrm{L}(\boldsymbol{\gamma}, \boldsymbol{\beta};\ \boldsymbol{Y})}{\mathrm{L}(\boldsymbol{\beta};\ \hat{\boldsymbol{\beta}}, \boldsymbol{\gamma})}.$$

The REML log-likelihood function is

$$\log \mathrm{L}(\boldsymbol{\gamma}; \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) = K - \frac{1}{2}\left[\log \mid \boldsymbol{V} \mid - \log \mid \boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X} \mid^{-1} + (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})\right].$$

Compare this with equation 2.3 and notice that the difference is the subtraction of the term

$$\log \mid \boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X} \mid^{-1}.$$

It is possible to obtain parameter estimates iteratively using the method of iterative generalized least squares (IGLS). IGLS gives the same results as ML but the details are not included here. However, the interested reader can refer to Applied Mixed Models in Medicine by Brown and Prescott [13].

2.2.2 Estimation of fixed effects

The fixed effects estimates are obtained by maximizing the likelihood by differentiating the log-likelihood with respect to $\boldsymbol{\beta}$ and set to zero. This gives the following formula

$$\boldsymbol{X}^T\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) = 0,$$

which leads to

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}. \tag{2.4}$$

The variance of $\hat{\boldsymbol{\beta}}$ is

$$\begin{aligned}\mathbf{Cov}(\hat{\boldsymbol{\beta}}) &= (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\mathrm{Cov}(\boldsymbol{Y})\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\\ &= (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{V}\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\\ &= (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}.\end{aligned}$$

2.2.3 Prediction of random effects

The random effects are treated as realizations from an underlying normal distribution with mean zero and a specified variance-covariance structure. Although their true values are unobserved, it is possible to obtain predictions of these effects based on the observed data. This is typically done using ML and REML. First, define a product of likelihoods in terms of $\boldsymbol{\beta}, \boldsymbol{b},$ and $\boldsymbol{\gamma}$

$$\mathrm{L}(\boldsymbol{\beta}, \boldsymbol{b}, \gamma; \boldsymbol{Y}) = \mathrm{L}(\boldsymbol{\beta}, \boldsymbol{\gamma}_{\boldsymbol{R}}; \boldsymbol{Y} \mid \boldsymbol{b})\mathrm{L}(\boldsymbol{\gamma}_{\boldsymbol{G}}; \boldsymbol{b})$$

where $\boldsymbol{\gamma_R}$ are the parameters in the $\mathbf{R}$ matrix and $\boldsymbol{\gamma_G}$ are the parameters in the $\mathbf{G}$ matrix.

Assume that $\boldsymbol{Y} \mid \boldsymbol{b}$ and $\boldsymbol{b}$ are multivariate normal distributed we get

$$\begin{aligned}\mathrm{L}(\boldsymbol{\beta}, \boldsymbol{b}, \gamma; \boldsymbol{Y}) \propto &\mid \boldsymbol{R} \mid^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Zb})^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Zb})\right) \\ &\times \mid \boldsymbol{G} \mid^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{b}^T \boldsymbol{G}^{-1}\boldsymbol{b}\right),\end{aligned}$$

and the log-likelihood is

$$\begin{aligned}\log(\mathrm{L}) = -&\frac{1}{2}[\log \mid \boldsymbol{R} \mid + (\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Zb})^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{Zb}) \\ &+ \log \mid \boldsymbol{G} \mid + \boldsymbol{b}^T \boldsymbol{G}^{-1}\boldsymbol{b}] + K.\end{aligned}$$

Now, by differentiating the log-likelihood with respect to $\boldsymbol{b}$ and equate to zero gives

$$\hat{\boldsymbol{b}} = (\boldsymbol{Z}^T \boldsymbol{R}^{-1} \boldsymbol{Z} + \boldsymbol{G}^{-1})^{-1} \boldsymbol{Z}^T \boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta}).$$

Using matrix manipulation and the fact that $\boldsymbol{V} = \boldsymbol{ZGZ^T} + \boldsymbol{R}$ the random effects estimates can be written in a more compact form

$$\hat{\boldsymbol{b}} = \boldsymbol{GZ}^T \boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}). \tag{2.5}$$

The variance of $\hat{\boldsymbol{b}}$ is

$$\mathbf{Cov}(\hat{\boldsymbol{b}}) = \boldsymbol{GZ}^T \boldsymbol{V}^{-1} \boldsymbol{ZG} - \boldsymbol{GZ}^T \boldsymbol{V}^{-1} \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{ZG}.$$

2.2.4 Henderson's mixed model equations

Another way to estimate $\boldsymbol{\beta}$ and $\mathbf{b}$ is through Henderson's mixed model equations (MME). From the MME equations, one can derive the best linear unbiased estimator (BLUE) of the fixed effects and the best linear unbiased predictor (BLUP) of the random effects. These estimates are the best estimates as long as the assumptions of the model are fulfilled. This procedure estimates $\boldsymbol{\beta}$ and $\mathbf{b}$ simultaneously without having to calculate $\boldsymbol{V}^{-1}$, as in large datasets, computing $\boldsymbol{V}^{-1}$ can be computationally expensive or numerically unstable. Henderson's equations avoid this by solving a linear system.

When $\mathbf{b}$ and $\mathbf{Y}$ are normally distributed we have

$$\begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{Y} \end{bmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mathbf{0} \\ \boldsymbol{X\beta} \end{pmatrix}, \begin{pmatrix} \boldsymbol{G} & \boldsymbol{GZ^T} \\ \boldsymbol{ZG} & \boldsymbol{V} \end{pmatrix}\right]$$

and maximizing the joint density function

$$\frac{\exp\left\{-\frac{1}{2}\left[\boldsymbol{b}^\top \mathbf{G}^{-1} \boldsymbol{b} + (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Zb})^\top \mathbf{R}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Zb})\right]\right\}}{(2\pi)^{N+q}\sqrt{\mid \mathbf{G} \mid\mid \mathbf{R} \mid}}$$

with respect to $\boldsymbol{\beta}$ and $\mathbf{b}$, Henderson's Mixed Model Equations becomes

$$\begin{bmatrix} \mathbf{X}^T\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^T\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{X} & \mathbf{G}^{-1} + \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{Z} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}^T\mathbf{R}^{-1}\mathbf{Y} \end{bmatrix}.$$

The solutions to Henderson's equations are the best linear unbiased estimator and predictor of $\boldsymbol{\beta}$ and $\mathbf{b}$.

$$\text{BLUE}(\boldsymbol{\beta}) = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y} = \hat{\boldsymbol{\beta}},$$

$$\text{BLUP}(\boldsymbol{b}) = \boldsymbol{G}\boldsymbol{Z}^T\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{b}}.$$

2.3 Joint mixed-effects models for longitudinal data

A joint mixed-effects model combines multiple outcomes or response variables into a single model, allowing for the simultaneous analysis of these outcomes while accounting for their potential correlation. If two outcomes are expected to change similarly over time, a joint model can capture that shared behaviour through correlated random effects. This means that a joint model can be more accurate in predictions since it uses information from both responses.

Consider home and hospital data as two data series. Let $Y_{ij}^{(k)}$ denote the outcome corresponding to measurement j for participant i and outcome type k. Note that $k \in \{1, 2\}$, where 1 defines home data and 2 hospital data and j is the number of spirometry measurements of the ith individual. $Y_{ij}^{(k)}$ can now be calculated as

$$Y_{ij}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} \cdot X_{ij} + b_{0i}^{(k)} + b_{1i}^{(k)} \cdot Z_{ij} + \varepsilon_{ij}^{(k)}$$

Here, Y_{ij} is the spirometry value of participant i at measurement j. Note that in a CF setting, X_{ij} and Z_{ij} coincides and represents the age of the ith participant at the jth measurement. Also, $(\beta_0^{(k)} + b_{0i}^{(k)})$ represent the intercept of participant i and $(\beta_1^{(k)} + b_{1i}^{(k)})$ the slope. Now, the response variable can be expressed as two-dimensional and allow for correlation between the two outcomes, home spirometry and hospital spirometry. See Appendix A for the specifications of the variance-covariance matrix of random effects $\mathbf{G}$ and residual variance-covariance matrix $\mathbf{R}$.

Combining $\boldsymbol{Y}_{ij}^{(0)}$ and $\boldsymbol{Y}_{ij}^{(1)}$ into a system of equations gives:

$$\begin{cases} \boldsymbol{Y}_i^{(1)} & = \boldsymbol{X}_i^{(1)}\boldsymbol{\beta}_i^{(1)} + \boldsymbol{Z}_i^{(1)}\boldsymbol{b}_i^{(1)} + \boldsymbol{\varepsilon}_i^{(1)} \\ \boldsymbol{Y}_i^{(2)} & = \boldsymbol{X}_i^{(2)}\boldsymbol{\beta}_i^{(2)} + \boldsymbol{Z}_i^{(2)}\boldsymbol{b}_i^{(2)} + \boldsymbol{\varepsilon}_i^{(2)}. \end{cases} \tag{2.6}$$

By stacking the two outcome vectors, the combined model can be represented as the joint system:

$$\boldsymbol{Y}_i = \begin{bmatrix} \boldsymbol{Y}_i^{(1)} \\ \boldsymbol{Y}_i^{(2)} \end{bmatrix}, \tag{2.7}$$

with corresponding block-diagonal design matrices and covariance structures for the fixed and random effects. This joint formulation enables the modeling of potential correlations between home and hospital measurements through a joint covariance structure in the random effects, which includes cross-covariances between the outcomes. Estimation of model parameters can then proceed as described in the preceding sections, using ML, REML, or IGLS as described previously.

The joint model enables estimation of shared individual-specific trends via a structured covariance matrix for the random-effects. Moreover, by modeling two longitudinal data series simultaneously, the joint model could use the information of correlation between the two outcomes, and perhaps, achieve better predictive accuracy in the outcome of interest (in our case, hospital spirometry).

In summary, to attain unbiased estimates REML is used. REML uses the BLUE estimates of the fixed-effects to eliminate them from the log-likelihood function so that it is only a function of the variance parameters. The estimation of the fixed effects, random effects and variance and covariance parameters are all built upon maximizing the log-likelihood function. The optimization is done using an iterative procedure, for example, the Newton-Raphson method which is the method that is most commonly used. First, the log-likelihood is maximized with respect to the variance parameters while treating the fixed effects as constants, where an initial value can be attained from ordinary least squares or SLR by assuming independence. Then, the estimates of the variance parameters are treated as fixed and maximization of the log-likelihood with respect to the fixed effects can be done. This is repeated until convergence.

2.4 Literature review

This chapter is concluded with a brief review of the literature on mixed-effects modeling and its applications in CF research.

Joint linear mixed-effects models are commonly used when analyzing multiple outcomes simultaneously. The outcomes can be of different types such as continuous, binary or survival outcomes. Ivanova [15] et al. introduced a random effects longitudinal model that uses different types of outcomes. Outcome variables does not always have to follow a normal distribution but they can be skewed or follow a different distribution which is the setting Huang et al. [14] discusses. This setting is applied on a joint nonlinear mixed-effects models. Furthermore, Li [16] presents a joint mixed-effects model with multivariate longitudinal data with continuous, proportion, count and binary responses and the association between the variables are captured through correlated random effects. Daza et al. [7] compares a joint model to a linear mixed-effects model in a randomized control trial context.

Advanced applications also focused on causal inference. [24] integrated joint modeling and g-computation to estimate causal effects in the presence of time-varying and unmeasured confounding. A Bayesian extension was proposed in [33], using generalized linear mixed models with posterior distributions to evaluate subgroup-specific treatment effects.

In terms of statistical estimation and implementation, [2] provided a comprehensive description of the `lmer` function in the `lme4` package for `R`. For researchers with less experience, introductory materials such as [6], [8], and [12] offered theoretical foundations and practical guidance on implementing linear mixed-effects models. Magezi [19] focused on within-participant designs and offered tools for users unfamiliar with `R`. The basic estimation principles of BLUE and BLUP with focues on their matrix form were covered in [23], while [21] discussed how iterative weighted least squares could be used to obtain maximum likelihood estimates. Deciding model structures in longitudinal data was addressed in [11], discussing balanced vs. unbalanced designs and missing data. Am election procedure of the fixed and random effects was discussed in [9]. Luke [17] discusses the procedure of evaluating the significance of fixed effects using p-values.

In CF research, the use of longitudinal data and mixed-effects models is essential for the understanding of disease progression. Mixed-effects models have been used to analyze repeated measures such as percent predicted FEV_1, with studies showing that modeling choices such as linear vs. non-linear or marginal vs. mixed-effects can affect estimates of lung function decline [28]. While home spirometry tends to yield lower and more variable values compared to supervised clinical tests [22, 1, 10], it remains a viable option for weekly monitoring [4]. Integrating home and hospital data in telehealth settings has shown outcomes comparable to traditional in-person care [20]. Longitudinal models also help answer questions about risk factors. Variability in lung function during stable periods at home has been found to exceed that seen in clinic settings, highlighting the need for cautious interpretation of home data [30]. Overall, the quality of CF trend assessment depends on methodological choices and clinical interpretation [25, 29].

In summary, linear mixed-effects models are well-established for analysing longitudinal clinical data, including applications in cystic fibrosis. These models have been used to study lung function decline over time, although their implementation varies in complexity. However, joint models have not been thoroughly evaluated in this context, particularly regarding the efficient use of home spirometry measurements to support individual-level assessments. In the remainder of this thesis, we develop and evaluate joint linear mixed-effects models for combined home and hospital spirometry data, using both simulation studies and real-world applications.

3

Simulation experiments

This chapter presents simulation experiments to evaluate the performance of joint linear mixed-effects models (JLMM) compared to univariate mixed-effects models (ULMM) and simple linear regression (SLR) in the context of modeling bivariate longitudinal data. The chapter begins by introducing the simulation objectives and data-generating process, followed by a description of the model specifications, simulation settings, and evaluation metrics. The final sections present the results of the experiments and summarise their implications for model selection in longitudinal analysis of spirometry data.

3.1 Introduction and objectives

Through simulations, we aimed to evaluate SLR, ULMMs and JLMMs performed in different settings and the simulations were designed to answer the following questions:

1. How does a joint mixed-effects model compare to a univariate mixed-effects model and simple linear regression in terms of estimation and prediction accuracy?
2. How is predictive performance affected by changes in the number of participants?
3. How does the prediction error vary with the correlation between subject-specific trends?
4. Can hospital spirometry measurements be replaced by additional home spirometry measurements without compromising accuracy?
5. How frequently do convergence issues occur, and how often is the estimated variance-covariance matrix of the random effects is singular?

These questions will be evaluated through simulation scenarios that mimic characteristics commonly observed in cystic fibrosis, including irregular measurement schedules, differing frequencies of home and hospital spirometry, and declining lung function over time. The simulations examine how model accuracy and stability are influenced by sample size, measurement frequency, and the correlation between subject-specific trends.

3.2 Simulation setup

This section describes the data-generating process used in the simulations, including the structure, parameter settings, and assumptions designed to reflect longitudinal spirometry data typical of individuals with cystic fibrosis. Unless otherwise stated, the number of participants included in the simulations was set to 100. The intention of the input parameter values were that they should produce data that was similar to a real CF data set. The parameters were selected to produce data that resembled CF data. First, hospital spirometry had higher values than home spirometry. Second, the regression lines for home spirometry were specified to have steeper slopes. The number of home measurements varied between individuals but averaged 12 per year. In contrast, each participant had 4 hospital spirometry measurements annually. Measurements were taken at different time points across participants. The random effects were assigned a variance of 1 around the fixed effects, and the residuals was also set to have a variance of 1.

3.2.1 Formulation of the model

A total of four fixed effects was considered and these were generated with the normal distribution using the `mvrnorm` function from the `MASS` package [31]. The fixed effects were set directly and the random effects were assumed to be normally distributed around the fixed effects. The fixed effects were set by taking into account of the fact that home data systematically exhibit lower values was accounted for. Now, from the simulated values we have values of the regression parameters of each individual and hence data can be generated by adding residuals to each regression model. For example, to generate data that mimics hospital spirometry at a certain age, the following was implemented:

$$Y_{ij}^{(2)} = (b_{0i}^{(2)} + \beta_{0i}^{(2)}) + (b_{1i}^{(2)} + \beta_{1i}^{(2)}) \cdot Age_{ij}^{(2)} + \varepsilon_{ij}^{(2)}.$$

Consider figure 3.1 which illustrates simulated data of four individuals. Each individual provides home and hospital measurements and the regression line of each measurement are displayed as gray lines. Note that, in reality, the differences between home and hospital measurements might not be as big as it appears in the figure. This is a toy data set and is provided as an illustration.

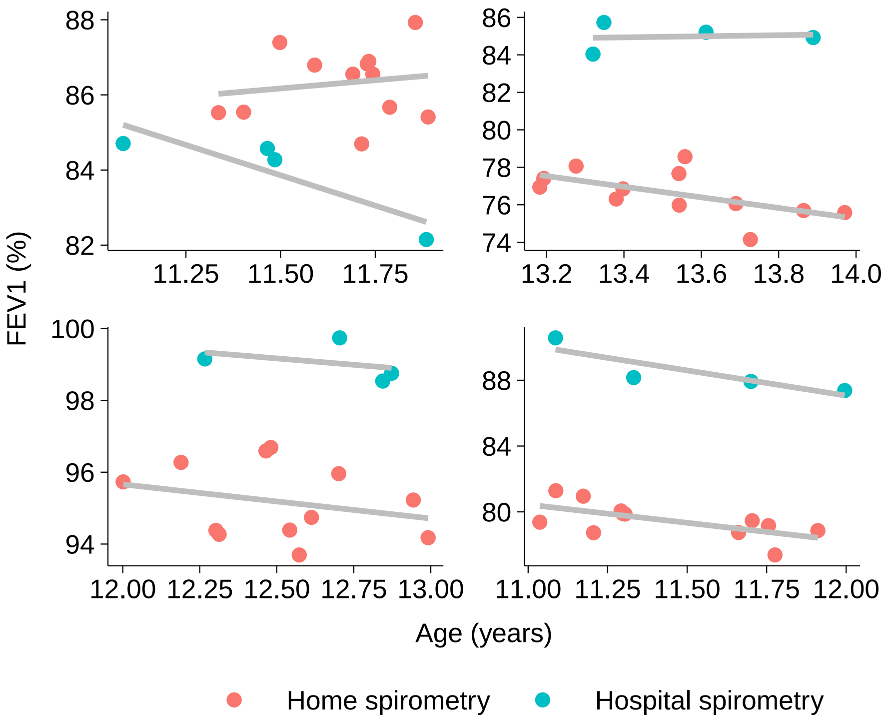

Figure 3.1: Simulated data of four participants. Including random regression line of home and hospital spirometry for each individual represented as the grey lines.

3.2.2 Specification of the covariance structure

The random effects were simulated using a multivariate normal distribution, with a variance-covariance matrix called **G**, which is a 4×4 matrix that includes the variances and correlation terms between the intercept and slope of both home and hospital data. **G** is symmetric and structured as:

$$\mathbf{G} = \begin{bmatrix} 1 & \rho & -\rho & 0 \\ \rho & 1 & 0 & -\rho \\ -\rho & 0 & 1 & \rho \\ 0 & -\rho & \rho & 1 \end{bmatrix}.$$

It should be noted that **G** corresponds to the variance-covariance matrix from theory. **G** has to be positive definite in order to generate values. To achieve this, the R function `nearPD` from the `Matrix` package [3] was used, which adjust the input matrix and returns the nearest symmetric positive definite matrix to the input matrix. Because of this adjustment, the matrix that was used in the simulations did not necessarily have the exact same covariance terms as specified in **G**. The columns in **G** are in the following order $b_0^{(1)}, b_0^{(2)}, b_1^{(1)}, b_1^{(2)}$. The choices of variance and covariance are arbitrary and can be chosen differently depending on the purpose of the simulation.

3.2.3 Simulation settings

The parameters varied in this thesis were the number of participants, correlation between random effects, number of home spirometry measurements, and time horizon. The number of participants were varied between 5 and 50 to simulate for smaller sample sizes. The correlation was varied between 0 and 0.95 as negative correlation were not of practical interest and correlation of 1 is practically not possible. The number of hospital measurements was fixed to four and home measurements varied between 4 to 15. Lastly, time horizons were either 1, 3, or 5 years.

3.2.4 Model specification

Table 3.1 presents a synthetic data set designed to illustrate and clarify the relationship between the data structure and the corresponding model specification. The model specifications of the joint LMM (JLMM), univariate LMM (ULMM) and SLR are provided in listing 1. The `lm` is a package used to fit linear models and was used to fit SLR. The `lme4` package was used to fit ULMM and JLMM. The data, i.e. `df`, used for JLMM corresponds to the data in Table 3.1. Note that in our context, for example the first FEV_1 value in Table 3.1 corresponds to $y_{11}^{(1)}$ in mathematical notation.

Table 3.1: Example data set mimicking cystic fibrosis (CF) spirometry measurements across two settings. The table shows FEV_1 values by age for two individuals, with measurements taken at home (Setting = 1) and in hospital (Setting = 2), corresponding to the joint model specification.

ID	Setting	Age	FEV_1
1	1	12	90
1	1	12.5	85
1	2	12.1	95
1	2	13	90
2	1	15	100
2	1	15.5	95
2	2	15.1	105
2	2	16	103

The model specification returned a total of four fixed-effects intercepts, corresponding to the population regression parameters of the home spirometry and hospital spirometry. Also, the model specification returned the four random effects values for each individual. The setting parameter was treated as a categorical variable, allowing different coefficients to be estimated for each setting. An interaction term was included to specify the effect of age on FEV_1 depending on the setting which were the random effects of the slope parameters. Note that the data used in each model specification were different. The data used in SLR includes only hospital spirometry from a specific individual, where the data for ULMM includes only hospital spirometry from all individuals and JLMM utilizes the full dataset (i.e., home and hospital data across all individuals).

```r
# Simple Linear Regression
lm(FEV1 ~ Age, data = df)

# Univariate linear mixed-effects model
lmer(
  FEV1 ~ Age + (1 + Age | ID),
  data = df
)

# Joint linear mixed-effects model
lmer(
  FEV1 ~ -1 + as.factor(Setting) + as.factor(Setting):Age +
         (-1 + as.factor(Setting) + as.factor(Setting):Age | ID),
  data = df
)
```

Listing 1: Model specifications for the joint linear mixed-effects model (JLMM), the univariate linear mixed-effects model (LMM), and simple linear regression (SLR), all implemented in `lmer` or `lm` in `R`.

3.2.5 Evaluation metrics

The performance of the joint model was compared to traditional models in terms of prediction accuracy. Additionally, the properties of the proposed methodology was empirically evaluated through simulations. To answer the previous listed questions, the following evaluation metrics was used to quantify prediction or estimation error:

1. Current value (CV), i.e., error in estimating the actual value at a given time point after removing the target observation: $(y_{i,j} - \hat{y}_{i,-j})$,
2. Current mean (CM) at a specific point in time, i.e., error in estimating the predicted value at a given time point after removing the target observation: $(\hat{y}_{i,j} - \hat{y}_{i,-j})$,
3. Subject-specific trend (SST), i.e., error in estimating the individual-specific trend after excluding the target observation: $(\hat{\beta}_1 + \hat{b}_{1i}) - (\hat{\beta}_{1i,-j} + \hat{b}_{1i,-j})$,
4. Population average trend (PAT), i.e., error in estimating the population trend after excluding the target observation: $\hat{\beta}_1 - \hat{\beta}_{1,-j}$.

Here, $y_{i,j}$ was the i:th participant value of the response at measurement j. $\hat{y}_{i,j}$ was the i:th participant predicted response at measurement j of the model that was fitted on the entire data set and $\hat{y}_{i,-j}$ was the predicted value at measurement j where the j:th hospital value was removed. $\hat{\beta}_{1i}$ and $\hat{b}_{1i}$ was the estimated fixed and random effects slope parameters, respectively, on the full data set. $\hat{\beta}_{1i,-j}$ and $\hat{b}_{i,-j}$ were the estimated fixed and random effects slope parameters, respectively, on the modified data set. $\hat{\beta}_1$ was the estimated fixed effects on the full data set and $\hat{\beta}_{1,-j}$ were the estimated fixed effects on the modified data. In a CF context, the response variable represents spirometry data and X_{ij} represent the age of participant i at measurement j.

Figure 3.2 illustrates the differences of current value, current mean, and subject-specific trend. Current value is defined as the difference between the thick black line and the black dot. Current mean is defined as the difference between the thick and the dashed lines. Subject-specific trend represents the difference between the parameter values of the thick line and dashed line. Note that, PAT is not visible in figure 3.2, because the population regression lines were omitted to avoid cluttering the image.

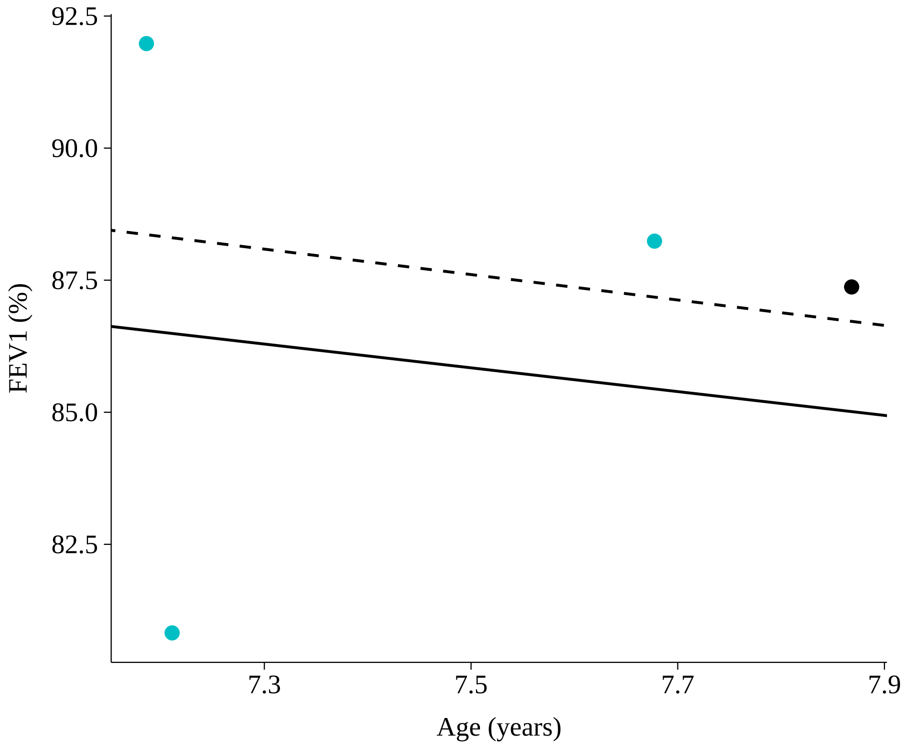

Figure 3.2: Hospital spirometry data of one individual represented by the blue and black dots. The black dot represents the removed spirometry measurement. The black line represent the individual hospital regression line on all data (i.e. blue and black dots). The black dashed line represent the individual regression line on blue dots only.

Performance will be assessed using the root mean square error (RMSE) for each of the differences defined above. The RMSE of the differences above provides evaluation metrics as a standardized measure of the average prediction error and was computed as the following:

$$\text{RMSE(CV)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y(t_{i,j}) - \hat{y}_{i,-j}(t_{i,j}))^2}, \tag{3.1}$$

$$\text{RMSE(CM)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}(t_{i,j}) - \hat{y}_{i,-j}(t_{i,j}))^2}, \tag{3.2}$$

$$\text{RMSE(SST)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((\hat{\beta}_1 + \hat{b}_{1i}) - (\hat{\beta}_{1i,-j} + \hat{b}_{1i,-j}))^2}, \tag{3.3}$$

$$\text{RMSE(PAT)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}_1 - \hat{\beta}_{1,-j})^2}. \tag{3.4}$$

To calculate RMSE, a variation of cross-validation was implemented. First, a synthetic data set was generated. Then, the last hospital measurement was removed for each individual. The original data, the original data with omitted hospital values, and only the omitted hospital values was saved as three data sets. A joint LMM was fitted on the original data and another one was fitted on the data with omitted hospital values. This was done using the `lmer` package. Using the estimated values from the model fit and extracting the correct values from the omitted data values, the RMSE of current value, current mean, subject-specific trend, and population average trend was calculated. A similar procedure was done for the ULMM, also using `lmer`, but only considering the hospital data. For SLR, the `lm` package was applied using only the hospital spirometry data for a given individual, thereby ignoring the parallel home measurement series.

Note that is possible that the estimated variance-covariance matrix of the random effects could have a variance or correlation term that is estimated to be zero. This is indicated by the `isSingular` function in the `lme4` package. In mathematical terms, the singularity issue arises when the estimated variance-covariance matrix lacks full rank.

3.3 Results

3.3.1 Overall performance across models

Figure 3.3 compares the performance of the joint mixed-effects model, the univariate mixed-effects model, and simple linear regression with the kernel density estimates of the root mean square error of the current value, current mean, subject-specific trend and the population average trend. The joint model and univariate model outperformed SLR in all scenarios. Visually, there was little to no difference in performance between the joint linear mixed-effects model and the univariate linear mixed-effects model for predicting the current value and the population-average trend. However, the univariate model performed better than the joint model on estimating the current mean and the subject-specific trend. Overall, the univariate model performed slightly better than the joint model. The difference between ULMM and JLMM are especially hard to visualize for current value across all time horizons. The difference in mean are: 0.003, 0.0009 and -0.001 in each case.

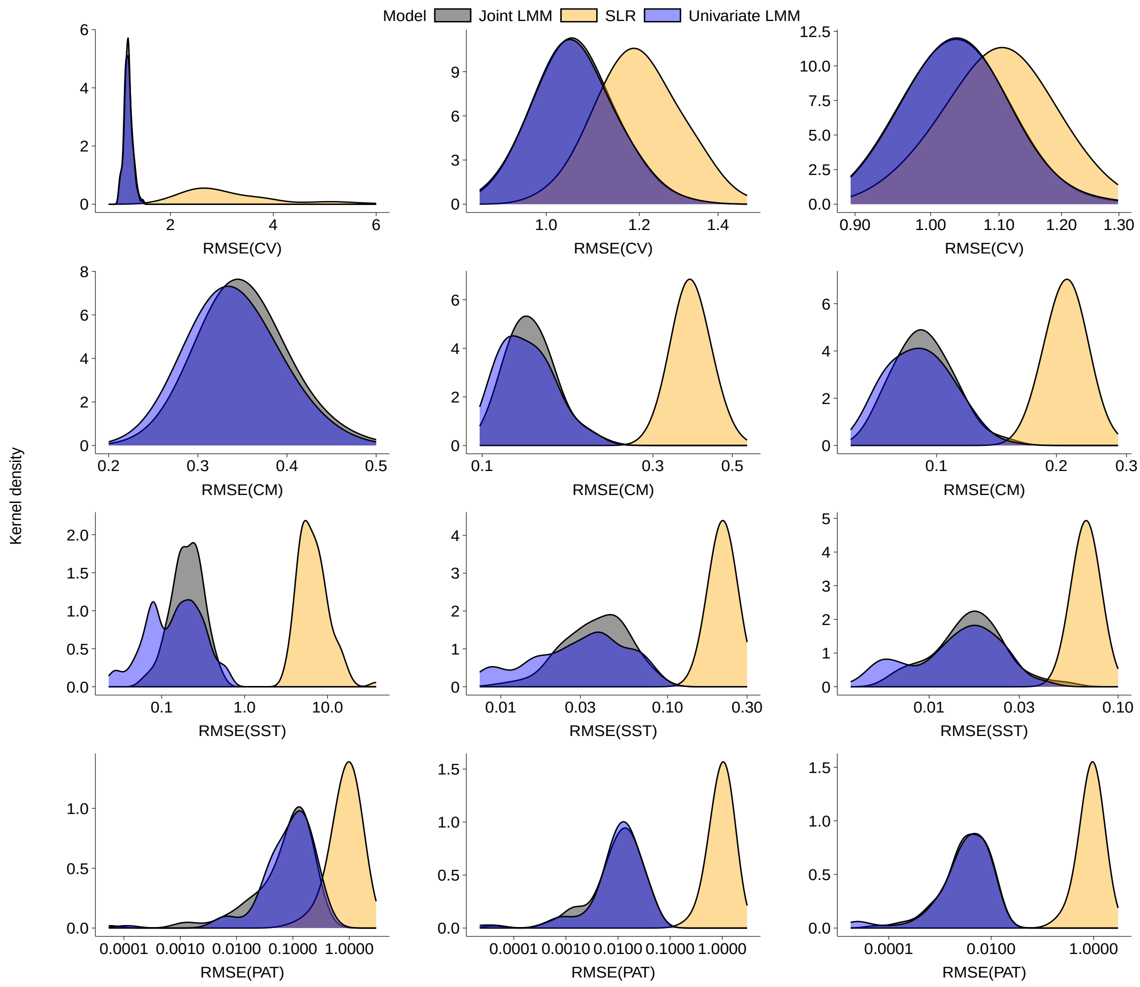

Figure 3.3: Kernel density estimates (KDEs) on the logarithm of RMSE of the current value, current mean, subject-specific trend and the population average trend across three time horizons: 1 year (left column), 3 years (middle column), and 5 years (right column). The rows correspond to the root mean square estimate distributions for four parameters: current value (fist row), current mean (second row), subject-specific trend (third row), and population average trend (fourth row).

3.3.2 Effect of sample size on performance

Figure 3.4 displays the RMSE of the current value, current mean, subject-specific trend and population average trend of lower values of sample size n. The univariate mixed-effects model and joint model approach similar performance as the number of individuals increased. However, the univariate model was slightly better in estimating the current value, current mean and subject-specific trend.

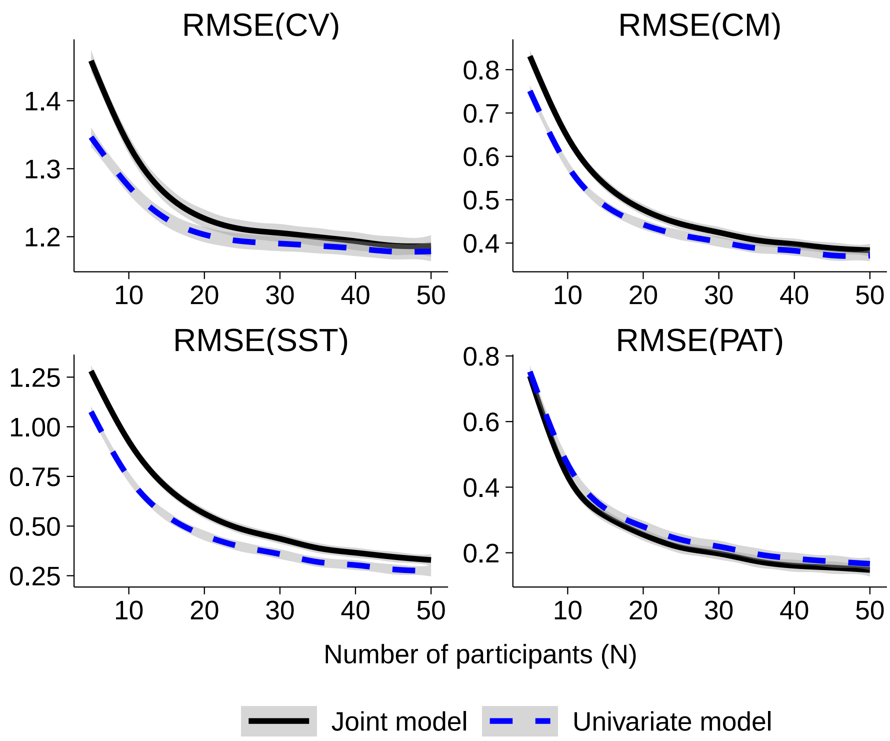

Figure 3.4: The estimated root mean square errors (RMSE) of current value (CV), current mean (CM), subject specific trend (SST), populations average trend (PAT) changing number of participants (N). The black line is the regression line of the joint LMM and the blue is of the univariate LMM. The shaded areas are 95% confidence intervals. The results are based on one year time horizon. The results of SLR are not included here because the results of SLR wont change as SLR depends only on the number of hospital measurements and not participants.

3.3.3 Impact of correlation between random effects

Figure 3.5 displays the percentage change in RMSE when comparing joint linear mixed-effects models to univariate linear mixed-effects models for predicting the current value, current mean, subject-specific trend, and population-average trend, as a function of the correlation between the random slope effects for home and hospital spirometry. This was done in order to quantify home much better or worse JLMM and ULMM were compared to each other at different correlation levels. The dashed horizontal line represents where JLMM and ULMM have an equal error. The RMSE of current value, current mean, subject-specific trend and population average trend of the JLMM decreases as the correlation increases. The JLMM had a lower error compared to the ULMM at a correlation of about 0.75 for the current value and current mean. The current value received a minor improvement of about 0.1 percent and the current mean around 4 percent. The ULMM was better for the subject-specific trend at all correlation levels while the opposite was true for the population average trend. The comparison with SLR is presented separately in Appendix B.

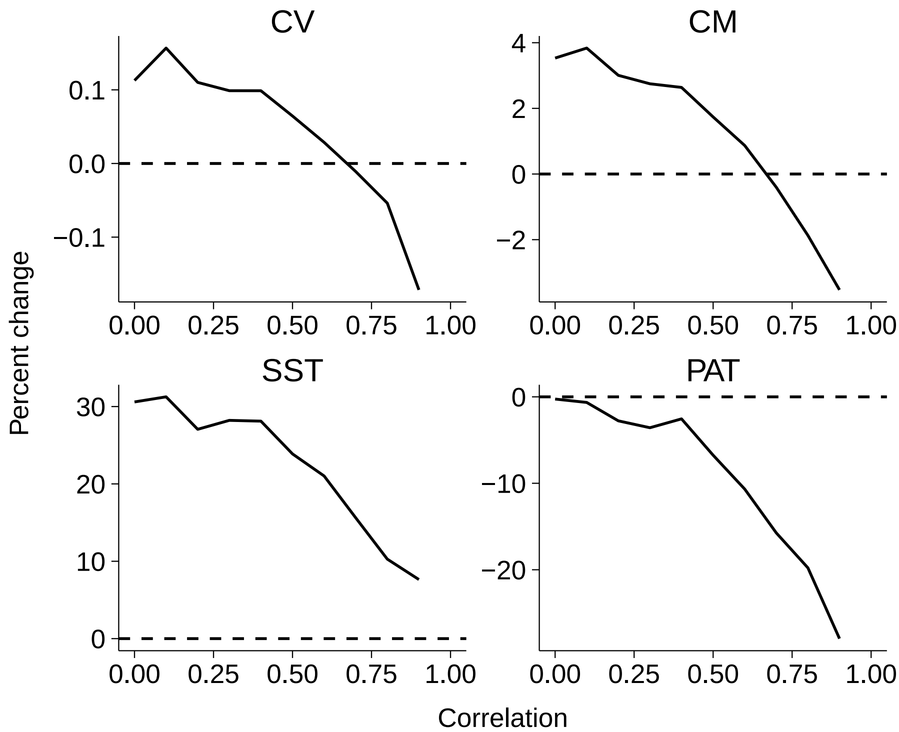

Figure 3.5: Percentage change in root mean square error (RMSE) when comparing joint linear mixed-effects models to univariate linear mixed-effects models for predicting the current value (CV), current mean (CM), subject-specific trend (SST), and population-average trend (PAT), as a function of the correlation between the random slope effects for home and hospital spirometry. The correlation between the home slope parameter and the hospital slope parameter have been varied. The dashed black lines represent where the joint LMM and univariate LMM are equal.

3.3.4 Substituting hospital with home spirometry

The results of the simulations when substituting a hospital spirometry with home spirometry are shown in figure 3.6. The joint model had a lower RMSE for the current value, current mean and the population average trend but for the subject-specific trend the univariate LMM had a lower error.

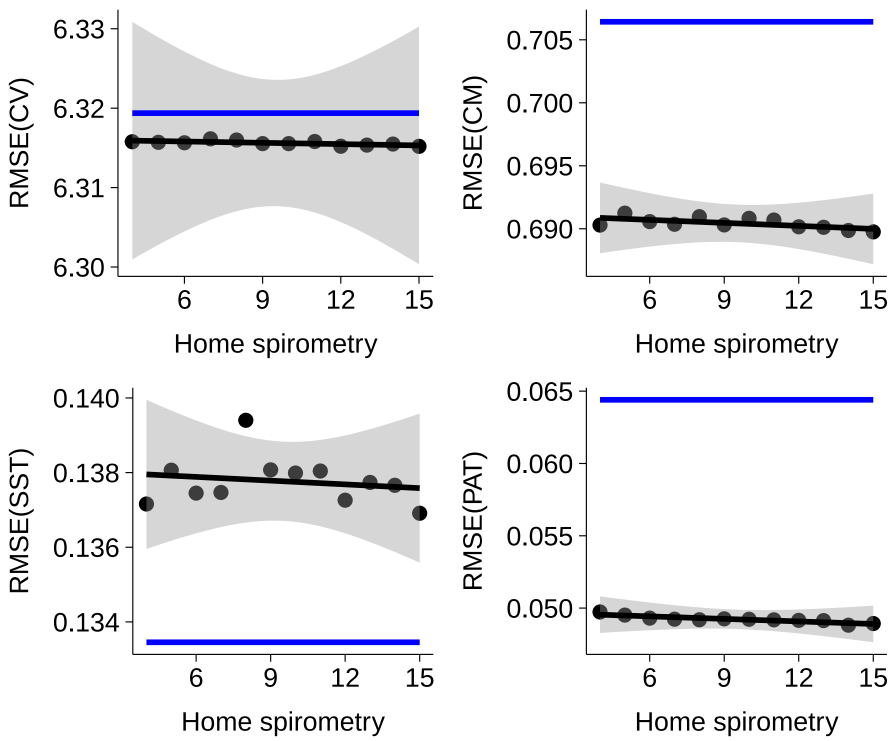

Figure 3.6: Root mean square errors (RMSE) for the current value (CV), current mean (CM), subject-specific trend (SST), and population-average trend (PAT) when one hospital measurement was replaced with varying numbers of home spirometry measurements ranging from 4 to 15. Dots represent mean RMSE values. Black lines indicate fitted joint LMM regression line with shaded 95% confidence intervals. The blue line represent the RMSE of the univariate LMM.

3.3.5 Model stability and the issue of singularity

Finally, the model's stability is evaluated by assessing the occurrence of singularities in the estimated random effects covariance structure. The proportion of simulated datasets yielding a singular covariance matrix estimate of the random effects is shown in figure 3.7. The JLMM had a high occurrence of singular variance-covariance matrices, especially for lower values on N. The ULMM had lowest occurrences of singular models and it approached to zero as the sample size N increased.

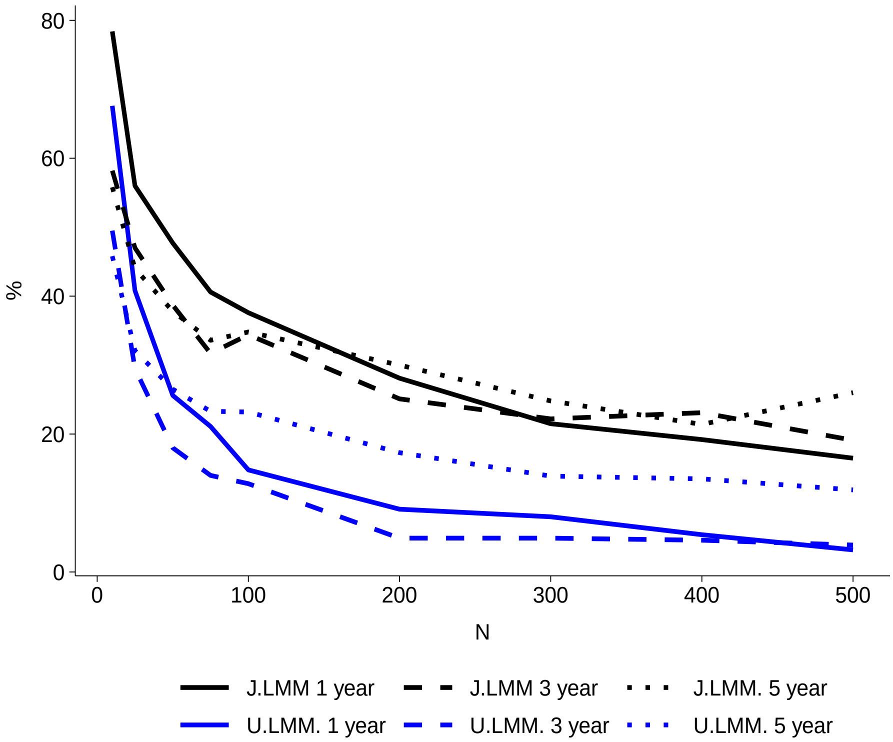

Figure 3.7: Proportion of simulated datasets producing a singular covariance matrix estimate of the random effects. Time horizons are: 1 year (thick line), three years (dashed line) and 5 years (dotted line). Black lines represent the joint model and blue lines the univariate model.

4
Application results

This chapter covers the implementation procedure applied to a real-world CF dataset. It begins by describing the dataset, followed by the methods and simulation setup. The final section presents the characteristics of the study population and summarizes the results.

4.1 Data

The dataset used in the application analysis comprises longitudinal spirometry measurements from individuals with cystic fibrosis (CF), collected at six specialised CF centers in Sweden between 2019 and 2024. The study population included both pediatric and adult patients, with the majority being children and adolescents. For each participant, two parallel time series of lung function measurements were available, i.e., home and hospital spirometry. In general, there were more home measurements available than hospital measurements, as hospital spirometry measurements were recommended to be four per year.

4.2 Methods

The objective of the application was to assess whether the most recent hospital spirometry measurement could be accurately predicted using a joint LMM, based on preceding hospital and home spirometry data. Specifically, the analysis compared the predictive performance of the JLMM to that of a ULMM and SLR, under the condition that the latest hospital measurement was excluded from the model fitting. The model specification was the same as in the simulations, see Listings 1.

Data cleaning involved the following steps. Spirometry values below 50 or above 150 were considered outliers and removed. Additional cleaning steps included removing missing values, discarding home spirometry more than three days apart from hospital measurements, and eliminating duplicate entries. Absolute spirometry values were converted to percent predicted values using the `rspiro` package, [18], and the global lung function initiative reference equations, based on each participant's demographic characteristics. The dataset was subdivided into paediatric and adult subpopulations to account for physiological and clinical differences in lung function trajectories, disease progression, and treatment response.

The lungs develop primarily the first years, then grows together with the rest of the body and this increases the lung volumes. The GLI equations, expressed as either a percentage or a z-score, is intended to account for the lung growth. Also, young individuals could respond differently to treatments compared to adults.

To calculate results on the real data the root mean square of the evaluation metrics: current value, current mean, subject-specific trend and population average trend was calculated (see Equations 3.1, 3.2, 3.3, and 3.4). The data was filtered to include only 1, 3, and 5 year time horizons in order to simulate how different time horizons affected the results. Also, a "thinning" procedure was done, i.e., twenty five and fifty percent of the hospital measurements were removed and the models were fitted on the data and compared to the original model fits. This was done to investigate how much data could be removed without loss of accuracy. Because of the thinning procedure, participants with four or fewer hospital spirometry measurements was removed as the univariate model became unidentifiable when some participant had only two data points.

4.3 Simulation setup

In order to provide mean values, confidence intervals, and p-values, a bootstrap analysis with size 100 was done on each population group, i.e., artificially creating 100 data sets. Each bootstrap sample included 100 participants and was modified to include the last 1, 3, and 5 years of data.

For each bootstrap resampled dataset, the predictive accuracy of JLMM, ULMM, and SLR was evaluated using four metrics: current value, current mean, subject-specific trend, and population average trend, as defined in Section 3.2.5. Due to the positively skewed distribution of the estimated root mean square error, all analyses were conducted on the logarithmic scale to stabilize variance and improve interpretability of model comparisons. Fold changes were computed to compare model performance for current value, current mean, subject-specific trend, and population average trend were calculated by first calculating the difference of the logarithm of the RMSE values between ULMM vs. SLR, JLMM vs. SLR, and JLMM vs. ULMM. Note, that in in the first two cases: SLR was used as benchmark and in the last case ULMM was used as benchmark. Then, a Welch's t-test was done on the log-transformed RMSE values and fold changes were calculated by exponentiating the mean differences of the logarithmic differences. The corresponding confidence intervals (95%) were calculated by exponentiating the confidence interval values of the t-test.

This means that a confidence interval that lies strictly below 1.0 indicates that the first method has a significantly lower error compared to the reference method, while a confidence interval strictly larger than 1.0 suggests significantly higher error. If JLMM vs. SLR yields a confidence interval below 1.0, it implies that the RMSE of JLMM is significantly lower than the RMSE of SLR.

4.4 Results

4.4.1 Characteristics of the study population

The raw data (i.e., before cleaning) included 162 individuals and 4,785 home spirometry measurements and 4,717 hospital measurements. Following the data cleaning process, which involved filtering out missing values and outliers, excluding home spirometry data, and removing duplicate entries, the number of participants in the pediatric group was 74 and there was a total of 1,857 home spirometry records and 1,486 hospital spirometry measurements in this group. Figure 4.1 explains the data cleaning and selection process as a flow chart and the statistical characteristics of the pediatric population are given in Table 4.1.

The youngest person was about 7 years old and the oldest was 18 years old where the ages was based on the latest age of each participant. The mean age of home spirometry was 13.9 years and 14.2 year respectively. The follow up times of home spirometry and hospital spirometry ranged between 0.1 to 4.7 and 1.1 to 5.8 respectively. The number of measurements for home and hospital spirometry ranged between 3 to 104 and 5 to 41. Figure 4.2 displays raw hospital based FEV_1 values (liters) across all participants.

FEV_1 increased with age during childhood and adolescence, peaked in early adulthood, and declined thereafter. Figure 4.3 and Figure 4.4 show the $FEV_1\%$ data of the pediatric sub-population and adult population respectively. Overall, there was a downward trend of spirometry measurements on both home measurements and hospital measurements.

Table 4.1: Characteristics of the study cohort for the pediatric population. The mean of age is based on the latest age of each participant.

	Home spirometry	Hospital spirometry
Number of measurements	1,857	1,486
Age (years)		
mean (SD)	13.9 (3.3)	14.2 (3.2)
range	7.0-18.0	7.1-17.9
Follow-up time* (years)		
median (IQR)	2.8 (1.5–4.0)	5.2 (3.8–5.6)
range	0.1–4.7	1.1–5.8
Number of measurements		
median (IQR)	19.0 (10.0–34.3)	20.0 (12.0–26.8)
range	3–104	5–41

*Time from the first to the last included FEV1-measurement.
IQR = interquartile range (25th and 75th percentiles).

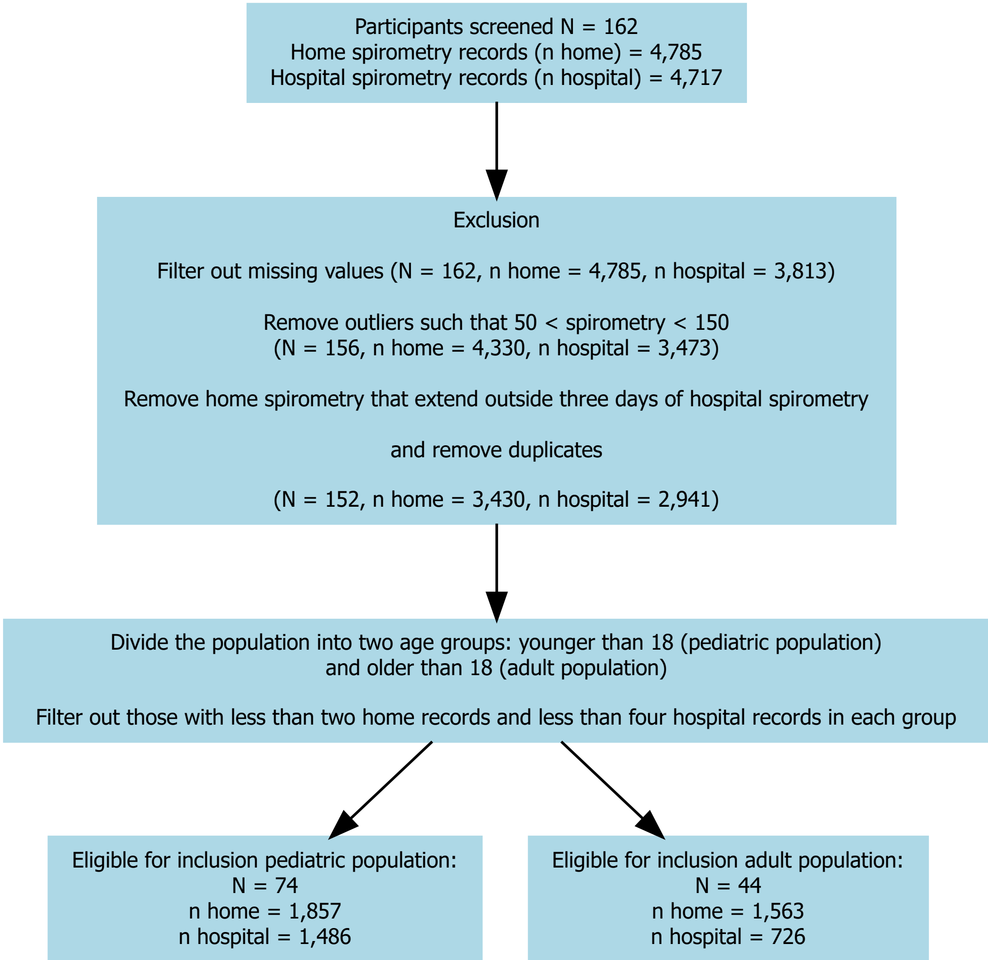

Figure 4.1: Overview of the data cleaning and selection process applied to the raw spirometry dataset, including exclusion criteria for outliers, missing values, duplicate records, and insufficient longitudinal data.

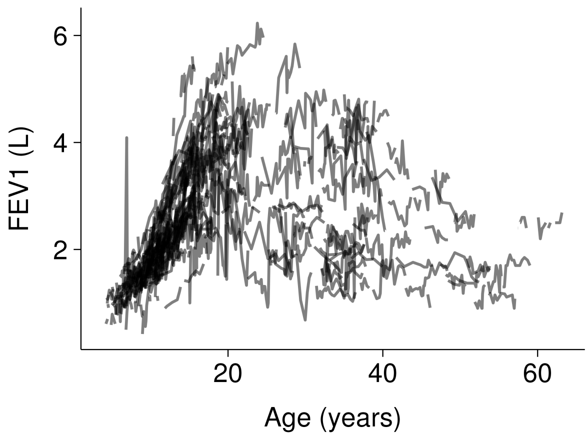

Figure 4.2: Hospital FEV_1 in liters (L) raw of data.

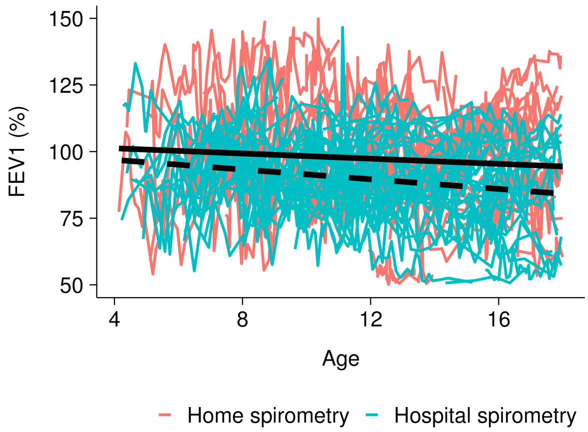

Figure 4.3: $FEV_1\%$ of age group 18 and younger including population regression lines. The dashed black line is the regression line of hospital spirometry and the black line is of the home spirometry. The regression lines of home and hospital spirometry records are displayed as a black line (home) and a dashed (hospital) line.

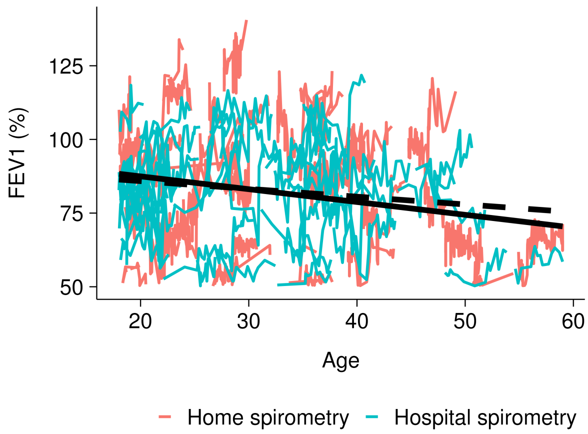

Figure 4.4: $FEV_1\%$ of age group 18 and older including population regression lines. The dashed black line is the regression line of hospital spirometry and the black line is of the home spirometry. The regression lines of home and hospital spirometry records are displayed as a black line (home) and a dashed (hospital) line.

4.4.2 Performance of JLMM, ULMM, and SLR

The results of the comparison of the JLMM, ULMM, and SLR of the RMSE of current value (CV), current mean (CM), subject-specific trend (SST), and population average trend (PAT) are presented in Table 4.2, Table 4.3, Table 4.4, and Table 4.5. The results are based on the pediatric population and the result for the adult population are found in Appendix B. Each table corresponds to current value, current mean, subject-specific trend and population average trend, respectively.

In the case of current mean, ULMM had a lower error than SLR for the remove-last scenario across all time horizons but SLR had lowest error in the 25% and 50% thinning-scenarios across all time horizons. The same can be said for JLMM vs. SLR where JLMM had a lower error in the remove-last scenario and SLR had lower error in the other scenarios. The RMSE between JLMM and ULMM for the current value was insignificant in the cases of: remove-last with three years time horizon, 50% thinning with one and five years time horizons.

Table 4.3 shows the results for the current mean. Here, ULMM and JLMM had a lower error than SLR in all scenarios. JLMM had lower error than ULMM in all scenarios except for the case of remove-last with one year time horizon, where the difference in error was insignificant.

The results for the subject-specific trend are shown in table 4.4 and here, the ULMM had a lower error compared to SLR in all cases, especially for one year time horizon. JLMM had a lower error compared to SLR in all cases. Finally, JLMM had a significantly lower error compared to ULMM in all scenarios except for three cases. First, in the 50%-thinning scenario, the difference was insignificant as the confidence interval includes 1.0. For the remove-last scenario with one year time horizon and for the 25%-thinning scenario, ULMM had a significantly lower error than JLMM.

The results of the population average trend are shown in table 4.5. Both ULMM and JLMM had a lower error compared to SLR in all cases. The difference in RMSE between JLMM and ULMM were insignificant in the cases of: remove-last scenario with one and three year time horizons. In the rest of the cases, JLMM had a significantly lower error than ULMM.

The results of the adult population were consistent with those observed in the pediatric cohort. SLR performed worse than ULMM in 31 out of 36 scenarios. SLR was worse than JLMM in 32 out of 36 scenarios. Finally, ULMM was worse than JLMM in 20 out of 36 scenarios.

Table 4.2: Fold change in root mean squared error (RMSE) for estimating the current hospital spirometry value (CV) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.83 (0.80, 0.83)	<0.001	0.81 (0.78, 0.84)	<0.001	0.97 (0.96, 0.98)	<0.001
	3 years	0.87 (0.86, 0.88)	<0.001	0.87 (0.86, 0.88)	<0.001	0.99 (0.98, 1.01)	0.124
	5 years	0.99 (0.98, 1.00)	0.003	0.95 (0.94, 0.95)	<0.001	0.96 (0.95, 0.97)	<0.001
25% thinning	1 year	2.01 (1.93, 2.08)	<0.001	1.96 (1.89, 2.04)	<0.001	0.98 (0.96, 0.99)	<0.001
	3 years	1.08 (1.06, 1.09)	<0.001	1.15 (1.13, 1.16)	<0.001	1.06 (1.05, 1.08)	<0.001
	5 years	1.09 (1.07, 1.10)	<0.001	1.07 (1.06, 1.09)	<0.001	0.99 (0.98, 1.00)	0.006
50% thinning	1 year	2.43 (2.29, 2.57)	<0.001	2.44 (2.30, 2.58)	<0.001	1.00 (0.99, 1.02)	0.523
	3 years	1.28 (1.26, 1.31)	<0.001	1.33 (1.31, 1.35)	<0.001	1.04 (1.01, 1.06)	<0.001
	5 years	1.16 (1.14, 1.18)	<0.001	1.16 (1.14, 1.17)	<0.001	1.00 (0.98, 1.01)	0.545

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

Table 4.3: Fold change in root mean squared error (RMSE) for estimating the current hospital spirometry mean (CM) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.26 (0.25, 0.27)	<0.001	0.25 (0.24, 0.26)	<0.001	0.98 (0.95, 1.01)	0.166
	3 years	0.58 (0.56, 0.60)	<0.001	0.44 (0.42, 0.45)	<0.001	0.75 (0.73, 0.78)	<0.001
	5 years	0.72 (0.71, 0.74)	<0.001	0.61 (0.59, 0.62)	<0.001	0.84 (0.82, 0.86)	<0.001
25% thinning	1 year	0.77 (0.74, 0.80)	<0.001	0.71 (0.68, 0.74)	<0.001	0.93 (0.90, 0.95)	<0.001
	3 years	0.76 (0.72, 0.81)	<0.001	0.62 (0.59, 0.64)	<0.001	0.81 (0.78, 0.84)	<0.001
	5 years	0.80 (0.78, 0.82)	<0.001	0.69 (0.67, 0.70)	<0.001	0.86 (0.85, 0.87)	<0.001
50% thinning	1 year	0.86 (0.83, 0.89)	<0.001	0.81 (0.78, 0.83)	<0.001	0.93 (0.92, 0.95)	<0.001
	3 years	0.89 (0.85, 0.92)	<0.001	0.66 (0.64, 0.68)	<0.001	0.74 (0.72, 0.76)	<0.001
	5 years	0.81 (0.79, 0.82)	<0.001	0.72 (0.70, 0.74)	<0.001	0.89 (0.87, 0.91)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

Table 4.4: Fold change in root mean squared error (RMSE) for estimating the subject-specific trend (SST) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.03 (0.03, 0.04)	<0.001	0.05 (0.04, 0.05)	<0.001	1.57 (1.31, 1.88)	<0.001
	3 years	0.46 (0.44, 0.48)	<0.001	0.28 (0.26, 0.30)	<0.001	0.62 (0.58, 0.66)	<0.001
	5 years	0.51 (0.50, 0.53)	<0.001	0.40 (0.38, 0.42)	<0.001	0.78 (0.75, 0.82)	<0.001
25% thinning	1 year	0.05 (0.04, 0.06)	<0.001	0.06 (0.06, 0.07)	<0.001	1.24 (1.07, 1.44)	0.006
	3 years	0.47 (0.43, 0.51)	<0.001	0.27 (0.24, 0.29)	<0.001	0.57 (0.53, 0.61)	<0.001
	5 years	0.47 (0.43, 0.50)	<0.001	0.38 (0.35, 0.40)	<0.001	0.81 (0.79, 0.83)	<0.001
50% thinning	1 year	0.04 (0.03, 0.05)	<0.001	0.05 (0.04, 0.05)	<0.001	1.14 (0.98, 1.33)	0.081
	3 years	0.47 (0.44, 0.50)	<0.001	0.25 (0.23, 0.26)	<0.001	0.53 (0.49, 0.56)	<0.001
	5 years	0.46 (0.43, 0.49)	<0.001	0.39 (0.36, 0.41)	<0.001	0.83 (0.81, 0.87)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

Table 4.5: Fold change in root mean squared error (RMSE) for estimating the population average trend (PAT) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.03 (0.02, 0.04)	<0.001	0.03 (0.02, 0.04)	<0.001	0.98 (0.76, 1.28)	0.905
	3 years	0.03 (0.02, 0.04)	<0.001	0.02 (0.02, 0.03)	<0.001	0.84 (0.67, 1.04)	0.108
	5 years	0.05 (0.04, 0.06)	<0.001	0.02 (0.02, 0.03)	<0.001	0.43 (0.35, 0.53)	<0.001
25% thinning	1 year	0.06 (0.05, 0.08)	<0.001	0.05 (0.04, 0.06)	<0.001	0.75 (0.62, 0.91)	0.003
	3 years	0.05 (0.04, 0.06)	<0.001	0.03 (0.02, 0.03)	<0.001	0.59 (0.49, 0.70)	<0.001
	5 years	0.03 (0.02, 0.04)	<0.001	0.02 (0.01, 0.02)	<0.001	0.65 (0.48, 0.87)	0.005
50% thinning	1 year	0.20 (0.16, 0.26)	<0.001	0.17 (0.13, 0.21)	<0.001	0.83 (0.72, 0.96)	0.012
	3 years	0.18 (0.16, 0.21)	<0.001	0.08 (0.06, 0.10)	<0.001	0.43 (0.35, 0.53)	<0.001
	5 years	0.06 (0.05, 0.08)	<0.001	0.03 (0.02, 0.04)	<0.001	0.49 (0.36, 0.66)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

5

Discussion

5.1 Main findings

This thesis evaluated the utility of joint linear mixed-effects models (JLMM) for analysing longitudinal spirometry data in cystic fibrosis, in comparison with univariate linear mixed-effects models (ULMM) and simple linear regression (SLR). Across the simulation scenarios, JLMMs and ULMMs demonstrated better predictive performance than SLR across all evaluation metrics. In the simulation study, the results between JLMM and ULMM were inconclusive.

When implementing SLR, ULMM, and JLMM on the application data, ULMM and JLMM had better predictive performance compared to SLR in all scenarios except for the current value in some cases. In contrast to the simulation results, JLMMs performed better than ULMMs in most application scenarios. Specifically, in 25 out of 36 cases — across all evaluation metrics, time horizons, and thinning conditions — the confidence interval for the performance difference between JLMM and ULMM was strictly below 1.0, indicating significantly lower prediction error for JLMMs. These results support that JLMMs provides an advantage over ULMMs in a CF setting when making predictions.

The primary limitation of JLMMs identified in this study was their tendency to estimate a singular variance-covariance matrix, where variance or covariance components were estimated to be zero, negative one or positive one.

5.2 Interpretation and explanations

The simulation study showed that ULMMs had a lower tendency in estimating a singular variance-covariance matrix, suggesting they may be a more robust modeling choice. If the quality of home spirometry measurements continues to improve with technological advancement, it may become feasible to treat home and hospital measurements as a single data source. In that case, the univariate model may be sufficient, and the added complexity of a joint model would be unnecessary.

The reasons for the singularity issue could be that the random-effects structure of the model was too complex for the data, insufficient variation or correlated random effects. While the model may still provide estimates, the variance or correlation components looses informational value and prediction accuracy decreases. In standard applications, simplifying the random-effects structure by removing redundant terms may resolve the issue. However, in the context of joint modeling, where the primary objective is to leverage the correlation structure between multiple outcomes (e.g., home and hospital spirometry), removing random effects defeats the purpose of the model. A singular fit in this setting therefore suggests that the joint model cannot fully utilize the cross-outcome correlation, limiting its predictive advantage over univariate models. As such, the observed singularities reflect a fundamental trade-off between model complexity and the ability to exploit shared variance across modalities in finite data settings.

Both JLMMs and ULMMs offered a gain in RMSE of CV, CM, SST and PAT when compared to SLR, as SLR uses only hospital spirometry of one individual it looses valuable information. LMMs uses more data at its disposal and therefore predictions improve. The joint model incorporates all available data, whereas the univariate model relies solely on hospital data across all individuals. Another argument against SLR is that to increase its prediction accuracy, an increase in hospital measurements is needed. However, the recommendations of the number of hospital visits are typically four per year, as increasing hospital visits adds to the burden of care for individuals with CF and their families.

Regarding the sample size, a reduction in sample size is generally expected to lead to worse predictions and an increase in sample size leads to improved predictions. However, at some point, the addition of more measurements may no longer enhance accuracy and instead lead to diminishing returns.

5.3 Clinical and practical implications

The monitoring of individual trends for people with CF lung disease is recommended to be done with univariate or joint linear mixed-effects models over simple linear regression. The results support the use of joint linear mixed-effects models when modeling home and hospital spirometry in parallel. The practical consequences includes reduced health burden for CF patients and improved telehealth, which is a desired outcome for most CF patients [20].

5.4 Methodological considerations and future work

Joint LMMs tended to be computationally intensive due to the larger number of parameters and more complex model structure, which can increase the risk of convergence issues. Moreover, joint LMMs typically assume normality of the random effects and this may not always be the case for real-world data. It is also difficult to test the normality assumption in practice. In theory, it is possible to use a different distribution. However, in this case one would have to implement different software and this has not been done in this thesis. Additionally, it was of interest to examine how the model's performance changed when the variation in home measurements was reduced but because of the singularity issue, simulating different variance levels resulted in incomprehensible results as the model estimated variances as zero.
Future research could explore extensions to non-linear or non-Gaussian settings, the integration of additional clinical and patient-reported outcomes, or implementing a Bayesian approach to account for small sample size or departures from normality assumptions.

5.5 Conclusion

In summary, joint linear mixed-effects models offered improved predictive performance compared to both simple linear regression and univariate linear mixed-effects models to make predictions in a CF setting.

Bibliography

[1] Rohan Anand, Rebecca McLeese, John Busby, Jonathan Stewart, William D.C. Man, Judy Bradley, and Mike Clarke, *Unsupervised home spirometry versus supervised clinic spirometry for respiratory disease: A systematic methodology review and meta-analysis*, European Respiratory Review **32** (2023), 220248.

[2] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker, *Fitting linear mixed-effects models using lme4*, Journal of Statistical Software **67** (2015), 1–48.

[3] Douglas Bates, Martin Maechler, and Mikael Jagan, *Matrix: Sparse and dense matrix classes and methods*, 2024, R package version 1.7-0.

[4] Fabien Beaufils, Raphaël Enaud, François Gallode, Grégory Boucher, Julie Macey, Patrick Berger, Michael Fayon, and Stéphanie Bui, *Adherence, reliability, and variability of home spirometry telemonitoring in cystic fibrosis*, Frontiers in pediatrics **11** (2023).

[5] Ariel Berlinski, Pamela Leisenring, Lauren Willis, and Sandra King, *Home spirometry in children with cystic fibrosis*, Bioengineering **10** (2023), 242.

[6] Violet A. Brown, *An introduction to linear mixed-effects modeling in R*, Advances in Methods and Practices in Psychological Science **4** (2021).

[7] Julian F. Daza, Aya A. Mitani, Shabbir M.H. Alibhai, Peter M. Smith, Erin D. Kennedy, Mark A. Shulman, Paul S. Myles, and Duminda N. Wijeysundera, *Joint models inform the longitudinal assessment of patient-reported outcomes in clinical trials: a simulation study and secondary analysis of the restrictive vs. liberal fluid therapy for major abdominal surgery (RELIEF) randomized controlled trial*, Journal of Clinical Epidemiology **176** (2024), 111553.

[8] Lisa M. DeBruine and Dale J. Barr, *Understanding mixed-effects models through data simulation*, Advances in Methods and Practices in Psychological Science **4** (2021).

[9] Maud Delattre and Marie Anne Poursat, *An iterative algorithm for joint covariate and random effect selection in mixed effects models*, International Journal of Biostatistics **16** (2020), 75–83.

[10] Claire Edmondson, Nicole Westrupp, Christopher Short, Paul Seddon, Catherine Olden, Colin Wallis, Malcolm Brodlie, Francis Baxter, Jonathan McCormick, Susan MacFarlane, Richard Brooker, Margaret Connon, Salim Ghayyda, Lesley Blaikie, Rebecca Thursfield, Lynsey Brown, April Price, Erin Fleischer, Daniel Hughes, Christine Donnelly, Mark Rosenthal, John Wallenburg, Keith Brownlee, Eric W.F.W. Alton, Andrew Bush, and Jane C. Davies, *Unsupervised home spirometry is not equivalent to supervised clinic spirometry in children and young people with cystic fibrosis: Results from the climb-cf study*, Pediatric Pulmonology **58** (2023), 2871–2880.

[11] Garrett M. Fitzmaurice and Caitlin Ravichandran, *A primer in longitudinal data analysis*, Circulation **118** (2008), 2005–2010.

[12] Xavier A. Harrison, Lynda Donaldson, Maria Eugenia Correa-Cano, Julian Evans, David N. Fisher, Cecily E.D. Goodwin, Beth S. Robinson, David J. Hodgson, and Richard Inger, *A brief introduction to mixed effects modelling and multi-model inference in ecology*, PeerJ **2018** (2018), e4794.

[13] Helen and Robin Brown Prescott, *Applied mixed models in medicine*, Wiley, 2014.

[14] Yangxin Huang, Jiaqing Chen, and Chunning Yan, *Mixed-effects joint models with skew-normal distribution for HIV dynamic response with missing and mismeasured time-varying covariate*, International Journal of Biostatistics **8** (2014).

[15] Anna Ivanova, Geert Molenberghs, and Geert Verbeke, *Mixed models approaches for joint modeling of different types of responses*, Journal of Biopharmaceutical Statistics **26** (2016), 601–618.

[16] Haocheng Li, Yukun Zhang, Raymond J. Carroll, Sarah Kozey Keadle, Joshua N. Sampson, and Charles E. Matthews, *A joint modeling and estimation method for multivariate longitudinal data with mixed types of responses to analyze physical activity data generated by accelerometers*, Statistics in Medicine **36** (2017), 4028.

[17] Steven G. Luke, *Evaluating significance in linear mixed-effects models in R*, Behavior Research Methods **49** (2017), 1494–1502.

[18] Theodore Lytras, *rspiro: Implementation of spirometry equations*, 2023, R package version 0.4.

[19] David A. Magezi, *Linear mixed-effects models for within-participant psychology experiments: An introductory tutorial and free, graphical user interface (lmmgui)*, Frontiers in Psychology **6** (2015), 110312.

[20] Jakob Medbo, Henrik Imberg, Christine Hansen, Christina Krantz, Isabelle de Monestrol, and Marcus Svedberg, *Telemedicine and home spirometry in cystic fibrosis: A prospective multicenter study*, Pediatric Pulmonology **59** (2024).

[21] J A Nelder and R W M Wedderburn, *Generalized linear models*, Source: Journal of the Royal Statistical Society. Series A (General) **135** (1972), 370–384.

[22] Alex Paynter, Umer Khan, Sonya L. Heltshe, Christopher H. Goss, Noah Lechtzin, and Nicole Mayer Hamblett, *A comparison of clinic and home spirometry as longtudinal outcomes in cystic fibrosis*, Journal of Cystic Fibrosis **21** (2022), 78–83.

[23] Shayle R Searle, *The matrix handling of BLUE and BLUP in the mixed linear model*, Linear Algebra and its Applications **264** (1995), 291–311.

[24] Michelle Shardell and Luigi Ferrucci, *Joint mixed-effects models for causal inference with longitudinal data*, Statistics in Medicine **37** (2017), 829.

[25] Sanja Stanojevic, David A. Kaminsky, Martin R. Miller, Bruce Thompson, Andrea Aliverti, Igor Barjaktarevic, Brendan G. Cooper, Bruce Culver, Eric Derom, Graham L. Hall, Teal S. Hallstrand, Joerg D. Leuppi, Neil MacIntyre, Meredith McCormack, Margaret Rosenfeld, and Erik R. Swenson, *Ers/ats technical standard on interpretive strategies for routine lung function tests*, European Respiratory Journal **60** (2022).

[26] Marcus Svedberg, Per Gustafsson, Harm Tiddens, Henrik Imberg, Aldina Pivodic, and Anders Lindblad, *Risk factors for progression of structural lung disease in school-age children with cystic fibrosis*, Journal of Cystic Fibrosis **19** (2020), 910–916.

[27] Marcus Svedberg, Henrik Imberg, Per Gustafsson, Mela Brink, Håkan Caisander, and Anders Lindblad, *Chest X-rays are less sensitive than multiple breath washout examinations when it comes to detecting early cystic fibrosis lung disease*, Acta Paediatrica **111** (2021), 1253–1260.

[28] Rhonda Szczesniak, Eleni Rosalina Andrinopoulou, Weiji Su, Pedro M. Afonso, Pierre Réegis Burgel, Elizabeth Cromwell, Emrah Gecili, Enas Ghulam, Christopher H. Goss, Nicole Mayer-Hamblett, Ruth H. Keogh, Theodore G. Liou, Bruce Marshall, Wayne J. Morgan, Joshua S. Ostrenga, David J. Pasta, Sanja Stanojevic, Claire Wainwright, Grace C. Zhou, Gabriela Fernandez, Aliza K. Fink, and Michael S. Schechter, *Lung function decline in cystic fibrosis: Impact of data availability and modeling strategies on clinical interpretations*, Annals of the American Thoracic Society **20** (2023), 958–968.

[29] Rhonda Szczesniak, Sonya L. Heltshe, Sanja Stanojevic, and Nicole Mayer-Hamblett, *Use of FEV1 in cystic fibrosis epidemiologic studies and clinical trials: A statistical perspective for the clinical researcher*, Journal of Cystic Fibrosis **16** (2017), 318–326.

[30] Christina S. Thornton, Amalia S. Magaret, Lisa A. Carmody, Linda M. Kalikin, Richard H. Simon, John J. LiPuma, and Lindsay J. Caverly, *Quantifying variation in home spirometry in people with cystic fibrosis during baseline health, and associations with clinical outcomes*, Journal of Cystic Fibrosis **23** (2024), 321–328.

[31] W. N. Venables and B. D. Ripley, *Modern applied statistics with S*, Springer, New York, 2002.

[32] Brady T. West, Kathleen B. Welch, and Andrzej T. Galecki, *Linear mixed models: A practical guide using statistical software*, 2 ed., Chapman and Hall/CRC, Boca Raton, FL, 2014.

[33] Yizhen Xu, Ji Soo Kim, Laura K. Hummers, Ami A. Shah, and Scott L. Zeger, *Causal inference using multivariate generalized linear mixed-effects models*, Biometrics **80** (2024), ujae100.

A

Matrix formulation of the joint linear mixed-effects model

A.1 Design matrices and model components

The following section provides the mathematical formulation of the joint linear mixed-effects model applied throughout this thesis and had the following expression:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \tag{A.1}$$

$$\boldsymbol{b}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{G}), \tag{A.2}$$

$$\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_i). \tag{A.3}$$

$\mathbf{Y}$ is a random vector of responses, $\mathbf{X}$ is an $n \times p$ design matrix of observed values of explanatory variables, $\mathbf{Z}$ is the $n \times q$ known matrix of random-effects, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed-effects, $\mathbf{b}$ is the $q \times 1$ vector of random-effects where $\mathrm{Cov}(\boldsymbol{b}) = \boldsymbol{G}$ and $\boldsymbol{b} \sim N(\mathbf{0}, \boldsymbol{G})$ accounting for subject specific variations, $\boldsymbol{\varepsilon}$ is the vector of residuals.

$$\boldsymbol{Y}_i = \begin{bmatrix} Y_{i1}^{(1)} & \cdots & Y_{in_i}^{(1)} & Y_{i1}^{(2)} & \cdots & Y_{im_i}^{(2)} \end{bmatrix}^T \tag{A.4}$$

n_i and m_i are the number of home and hospital spirometry measurements, respectively.

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 0 & X_{i1}^{(1)} & 0 \\ \vdots & & \vdots & \vdots \\ 1 & 0 & X_{in_i}^{(1)} & 0 \\ 0 & 1 & 0 & X_{i1}^{(2)} \\ \vdots & & \vdots & \vdots \\ 0 & 1 & 0 & X_{im_i}^{(2)} \end{pmatrix} \tag{A.5}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0^{(1)} & \beta_0^{(2)} & \beta_1^{(1)} & \beta_1^{(2)} \end{bmatrix} \tag{A.6}$$

$$\boldsymbol{Z}_i = \begin{pmatrix} 1 & 0 & Z_{i1}^{(1)} & 0 \\ \vdots & & \vdots & \vdots \\ 1 & 0 & Z_{in_i}^{(1)} & 0 \\ 0 & 1 & 0 & Z_{i1}^{(2)} \\ \vdots & & \vdots & \vdots \\ 0 & 1 & 0 & Z_{im_i}^{(2)} \end{pmatrix} \tag{A.7}$$

$$\boldsymbol{b}_i = \begin{bmatrix} b_{i0}^{(1)} & b_{i0}^{(2)} & b_{i1}^{(1)} & b_{i1}^{(2)} \end{bmatrix} \tag{A.8}$$

$$\boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i1}^{(1)} & \vdots & \varepsilon_{in_i}^{(1)} & \varepsilon_{i1}^{(2)} & \vdots & \varepsilon_{im_i}^{(2)} & \end{bmatrix} \tag{A.9}$$

Note that $\boldsymbol{X}_i$ equals $\boldsymbol{Z}_i$ in this specification, as both fixed and random effects include only an intercept and an age-related slope. However, in general, the design matrices $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ may differ and include additional covariates, where $\boldsymbol{X}_i$ captures population-level (fixed effect) predictors and $\boldsymbol{Z}_i$ encodes subject-specific (random effect) predictors.

A.2 Covariance structure of the G matrix

The joint mixed-effects model has four random effects, i.e., the intercept and slope of home data and intercept and slope of hospital data. Note that this specification is not applied in general and alternative specifications may be more suitable in different contexts. Hence, the $\boldsymbol{D}$-matrix have the following expression:

$$\boldsymbol{G} = Var(\boldsymbol{b}_i) = \begin{pmatrix} \sigma^2_{b_0^{(1)}} & \sigma_{b_0^{(1)}b_0^{(2)}} & \sigma_{b_0^{(1)}b_1^{(1)}} & \sigma_{b_0^{(1)}b_1^{(2)}} \\ \sigma_{b_0^{(1)}b_0^{(2)}} & \sigma^2_{b_0^{(2)}} & \sigma_{b_0^{(2)}b_0^{(2)}} & \sigma_{b_0^{(2)}b_1^{(2)}} \\ \sigma_{b_0^{(1)}b_1^{(1)}} & \sigma_{b_0^{(2)}b_0^{(2)}} & \sigma^2_{b_1^{(1)}} & \sigma_{b_1^{(1)}b_1^{(2)}} \\ \sigma_{b_0^{(1)}b_1^{(2)}} & \sigma_{b_0^{(2)}b_1^{(2)}} & \sigma_{b_1^{(1)}b_1^{(2)}} & \sigma^2_{b_1^{(2)}} \end{pmatrix}. \tag{A.10}$$

A.3 Covariance structure of the R matrix

In this thesis, we assume a diagonal covariance matrix of the error terms, i.e.,

$$\boldsymbol{R}_i = \mathbf{Cov}(\varepsilon_i) = \sigma^2 I = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix} \tag{A.11}$$

$$\boldsymbol{\gamma}_{\boldsymbol{R_i}} = (\sigma^2) \tag{A.12}$$

This is the only version of $\boldsymbol{R}_i$ we use in this thesis but one should note that the model can be extended to include correlation between the residuals. If that is the case, $\boldsymbol{R}_i$ will no longer be a diagonal matrix. For example, one could use the Matérn covariance function or a linear exponent correlation structure. It is also possible to allow for different variance terms depending on the different measurements i.e. home spirometry and hospital spirometry.

B

Supplemental results

B.1 Additional simulation results

This supplemental section includes the results of RMSE for different correlation levels between the home and hospital slope parameter as seen in figure B.1. Additionally, table B.1 provides the characteristics of the study cohort for the adult population, while table B.2, B.3, B.4, and B.5 presents the application results for the adult population.

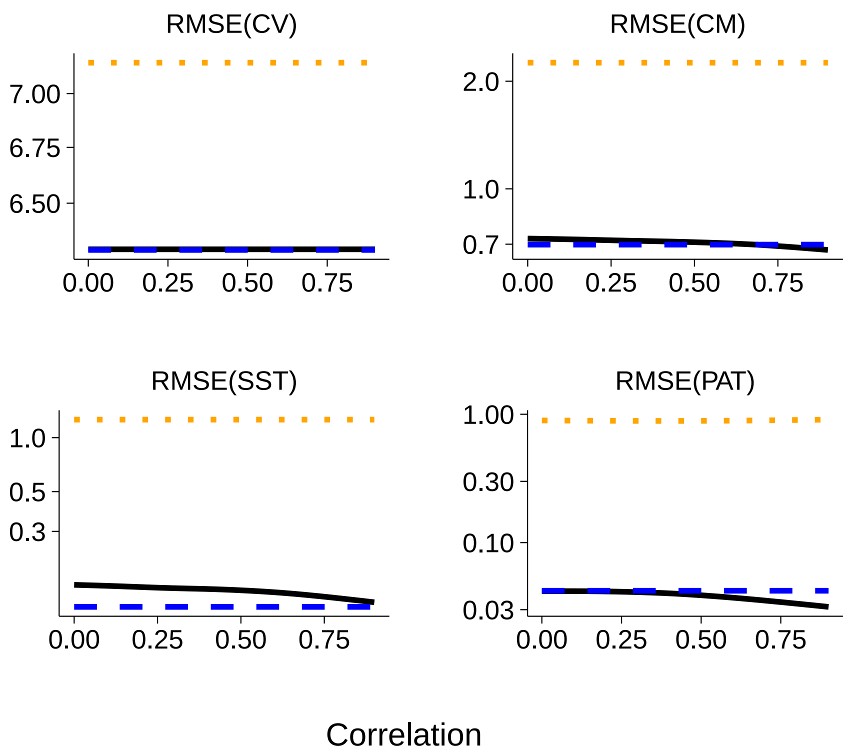

Figure B.1: Root mean square error for different correlation levels between the home and hospital slope parameter. The black line represent the joint LMM, blue dashed line represent univariate LMM and dotted orange is results of SLR.

B.2 Additional application results

Table B.1 shows the data statistic of the adult population. Table B.2, Table B.3, Table B.4, Table B.5 shows the results of the adult population.

Table B.1: Characteristics of the study cohort for the adult population. The mean of age is based on the latest age of each participant.

	Home spirometry	Hospital spirometry
Number of measurements	1,563	726
Age (years)		
mean (SD)	32.4 (10.6)	32.9 (10.5)
range	18.7–59.0	19.4–59.0
Follow-up time* (years)		
median (IQR)	2.4 (1.5–3.9)	5.2 (4.3–5.5)
range	0.3–7.6	1.4–5.7
Number of measurements		
median (IQR)	21.5 (10.5–54.5)	14.5 (12–22)
range	2–194	5–34

*Time from the first to the last included FEV1-measurement.
IQR = interquartile range 25th and 75th percentiles.

Table B.2: Fold change in RMSE for predicting the current hospital spirometry value (CV) using joint linear mixed models (JLMM), univariate linear mixed models (ULMM), and simple linear regression (SLR), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.62 (0.59, 0.64)	<0.001	0.70 (0.67, 0.73)	<0.001	1.13 (1.11, 1.15)	<0.001
	3 years	0.77 (0.75, 0.80)	<0.001	0.77 (0.75, 0.78)	<0.001	0.99 (0.97, 1.02)	0.454
	5 years	0.83 (0.82, 0.84)	<0.001	0.80 (0.79, 0.81)	<0.001	0.97 (0.95, 0.98)	<0.001
25% thinning	1 year	1.51 (1.45, 1.58)	<0.001	1.61 (1.54, 1.69)	<0.001	1.07 (1.05, 1.08)	<0.001
	3 years	1.18 (1.14, 1.23)	<0.001	1.08 (1.05, 1.11)	<0.001	0.91 (0.89, 0.93)	<0.001
	5 years	0.93 (0.91, 0.95)	<0.001	0.90 (0.89, 0.92)	<0.001	0.97 (0.96, 0.99)	<0.001
50% thinning	1 year	2.28 (2.16, 2.41)	<0.001	2.41 (2.28, 2.55)	<0.001	1.06 (1.04, 1.07)	<0.001
	3 years	1.46 (1.40, 1.51)	<0.001	1.27 (1.23, 1.30)	<0.001	0.87 (0.85, 0.89)	<0.001
	5 years	1.02 (0.99, 1.05)	0.215	0.96 (0.94, 0.98)	<0.001	0.94 (0.92, 0.96)	<0.001

Note. Values represent fold changes with 95% confidence intervals (CIs) derived from log-transformed RMSE. P-values are from Welch's t-tests on the log-transformed values.

Table B.3: Fold change in root mean squared error (RMSE) for estimating the current hospital spirometry mean (CM) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.22 (0.21, 0.23)	<0.001	0.30 (0.28, 0.32)	<0.001	1.35 (1.28, 1.42)	<0.001
	3 years	0.33 (0.31, 0.36)	<0.001	0.38 (0.36, 0.40)	<0.001	1.13 (1.04, 1.24)	0.007
	5 years	0.50 (0.48, 0.52)	<0.001	0.47 (0.45, 0.50)	<0.001	0.95 (0.91, 0.98)	0.001
25% thinning	1 year	0.82 (0.77, 0.86)	<0.001	0.79 (0.75, 0.83)	<0.001	0.97 (0.94, 1.00)	0.056
	3 years	0.93 (0.88, 0.99)	0.016	0.76 (0.72, 0.79)	<0.001	0.81 (0.77, 0.85)	<0.001
	5 years	0.87 (0.83, 0.91)	<0.001	0.78 (0.75, 0.81)	<0.001	0.90 (0.87, 0.93)	<0.001
50% thinning	1 year	0.85 (0.82, 0.88)	<0.001	0.83 (0.80, 0.86)	<0.001	0.98 (0.96, 1.00)	<0.001
	3 years	0.90 (0.86, 0.94)	<0.001	0.77 (0.74, 0.80)	<0.001	0.85 (0.82, 0.89)	<0.001
	5 years	0.95 (0.91, 1.00)	0.073	0.79 (0.76, 0.83)	<0.001	0.83 (0.80, 0.86)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

Table B.4: Fold change in root mean squared error (RMSE) for estimating the subject-specific trend (SST) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.03 (0.03, 0.04)	<0.001	0.10 (0.09, 0.12)	<0.001	3.28 (2.65, 4.06)	<0.001
	3 years	0.14 (0.12, 0.17)	<0.001	0.23 (0.21, 0.26)	<0.001	1.65 (1.34, 2.04)	<0.001
	5 years	0.24 (0.21, 0.26)	<0.001	0.24 (0.22, 0.27)	<0.001	1.02 (0.95, 1.09)	0.530
25% thinning	1 year	0.03 (0.02, 0.04)	<0.001	0.09 (0.07, 0.11)	<0.001	2.94 (2.25, 3.84)	<0.001
	3 years	0.23 (0.20, 0.28)	<0.001	0.33 (0.31, 0.36)	<0.001	1.43 (1.19, 1.73)	<0.001
	5 years	0.45 (0.40, 0.51)	<0.001	0.40 (0.36, 0.44)	<0.001	0.87 (0.80, 0.96)	0.003
50% thinning	1 year	0.02 (0.02, 0.02)	<0.001	0.05 (0.04, 0.06)	<0.001	2.60 (2.10, 3.21)	<0.001
	3 years	0.17 (0.14, 0.20)	<0.001	0.28 (0.26, 0.30)	<0.001	1.65 (1.38, 1.97)	<0.001
	5 years	0.47 (0.42, 0.53)	<0.001	0.38 (0.35, 0.42)	<0.001	0.81 (0.76, 0.88)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

Table B.5: Fold change in root mean squared error (RMSE) for estimating the population average trend (PAT) using univariate linear mixed models (ULMM) and joint linear mixed models (JLMM) incorporating home spirometry data, compared to simple linear regression (SLR). Values represent fold changes with 95% confidence intervals (CI), based on 100 bootstrap replications.

Scenario	Time horizon	ULMM vs SLR	*P*	JLMM vs SLR	*P*	JLMM vs ULMM	*P*
Remove last	1 year	0.08 (0.05, 0.11)	<0.001	0.22 (0.16, 0.30)	<0.001	2.76 (2.07, 3.66)	<0.001
	3 years	0.05 (0.04, 0.06)	<0.001	0.03 (0.03, 0.04)	<0.001	0.67 (0.46, 0.97)	0.035
	5 years	0.03 (0.02, 0.04)	<0.001	0.02 (0.02, 0.03)	<0.001	0.75 (0.57, 0.99)	0.039
25% thinning	1 year	0.07 (0.05, 0.11)	<0.001	0.11 (0.08, 0.02)	<0.001	1.56 (1.06, 2.29)	0.024
	3 years	0.09 (0.07, 0.11)	<0.001	0.03 (0.03, 0.04)	<0.001	0.36 (0.28, 0.48)	<0.001
	5 years	0.05 (0.04, 0.07)	<0.001	0.03 (0.02, 0.03)	<0.001	0.48 (0.38, 0.61)	<0.001
50% thinning	1 year	0.06 (0.04, 0.09)	<0.001	0.10 (0.08, 0.14)	<0.001	1.70 (1.31, 2.21)	<0.001
	3 years	0.22 (0.20, 0.26)	<0.001	0.09 (0.07, 0.11)	<0.001	0.40 (0.33, 0.49)	<0.001
	5 years	0.17 (0.15, 0.20)	<0.001	0.06 (0.05, 0.07)	<0.001	0.35 (0.29, 0.43)	<0.001

Note. Fold changes and 95% CIs were derived by exponentiating the mean difference on the log-transformed RMSE scale. Statistical comparisons were performed using Welch's t-test on the log-transformed values.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden
www.gu.se

UNIVERSITY OF
GOTHENBURG