Freie Universität Berlin

Modelling the distribution of health related quality of life of advanced melanoma patients in a longitudinal multi-centre clinical trial using M-quantile random effects regression

Riccardo Borgoni
Paola Del Bianco
Nicola Salvati
Timo Schmid
Nikos Tzavidis

School of Business & Economics
Discussion Paper

Economics

2015/19

Modelling the distribution of health related quality of life of advanced melanoma patients in a longitudinal multi-centre clinical trial using M-quantile random effects regression

Riccardo Borgoni* Paola Del Bianco† Nicola Salvati‡
Timo Schmid § Nikos Tzavidis¶

Abstract

Health-related quality of life assessment is important in the clinical evaluation of patients with metastatic disease that may offer useful information in understanding the clinical effectiveness of a treatment. To assess if a set of explicative variables impacts on the health-related quality of life, regression models are routinely adopted. However, the interest of researchers may be focussed on modelling other parts (e.g. quantiles) of this conditional distribution. In this paper we present an approach based on M-quantile regression to achieve this goal. We applied the proposed methodology to a prospective, randomized, multi-centre clinical trial. In order to take into account the hierarchical nature of the data we extended the M-quantile regression model to a three-level random effects specification and estimated it by maximum likelihood.

Keywords: Hierarchical data; influence function; robust estimation; quantile regression; multilevel modelling

1 Introduction

Assessing the health-related quality of life (HRQOL) forms an important part in the clinical evaluation of patients suffering from a metastatic disease and offers useful information for understanding the clinical effectiveness of a therapeutic option. For this reason

*Department of Economics, Management and Statistics, University of Milano - Bicocca, Milano, Italy, riccardo.borgoni@unimib.it

†Clinical Trials and Biostatistics Unit, Istituto Oncologico Veneto IOV - IRCCS, Padova Italy, paola.delbianco@ioveneto.it

‡Dipartimento di Economia e Management, Università di Pisa, Via Ridolfi, 10 - I56124 Pisa, Italy, salvati@ec.unipi.it

§Institute of Statistics and Econometrics, Freie Universität Berlin, Garystr. 21, Berlin, D-14195, Germany, Timo.Schmid@fu-berlin.de

¶Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ, UK, n.tzavidis@soton.ac.uk

HRQOL is nowadays considered as an important endpoint in oncological studies, helping physicians and patients to better understand the treatment outcomes, balancing among intent to cure, survival, side effects and quality of life, and to make appropriate decisions. HRQOL is a subjective, dynamic and multidimensional measure that includes physical, psychological and social domains. Its measurement evaluates the overall clinical benefit that a particular treatment offers to a patient.

HRQOL is evaluated through self assessment questionnaires measuring various aspects of quality of life. Data is frequently collected at various time points for investigating how the disease and treatment impact upon an individual's well-being over time.

Many clinical trials also enrol patients from different medical centres since conducting a multi-centre trial enhances the generalisation of findings, allowing researchers to evaluate the efficacy of a therapy in a variety of patients and settings, and makes it possible to investigate the effect of treatments when it is difficult, or even impossible, for a single centre to recruit the required number of patients. However, multi-centre studies involve an inter-centre variability as a result of differences in applying the study protocol procedures (Localio et al., 2001; Guthrie et al., 2012). Hence, a relevant issue in HRQOL is centre heterogeneity. The clustered structure of a multi-centre trial is neither accidental nor ignorable and ignoring it may lead to erroneous statistical inference.

Studies of multi-centre clinical trials quantify centre differences by using either a fixed effect or a random effect approach. In clinical trials with a relatively large number of centres each with a small number of patients the estimated centre-specific treatment effects may be unstable and the loss of efficiency can be severe (Sakamoto et al., 1999; Yamaguchi et al., 2002).

Centre effects can therefore be more conveniently specified by using a multilevel (random effects) modelling approach (Goldstein, 2003) where patients are hierarchically clustered within centres. Methods to design and analyse HRQOL studies, including the use of random effect models, are thoroughly discussed by Fairclough (2002).

Typically, a regression model offers a summary of the expected value of the conditional distribution of the outcome variable given the set of explanatory variables. Researchers, however, may be interested in modelling other parts (e.g. quantiles) of this conditional distribution. This may be the case if one is interested for example in understanding whether a treatment has a differential impact on patients with different HRQOL status. For instance, patients with a poor HRQOL status might suffer more when receiving one of the treatments under evaluation, than those who are in a better HRQOL state, as a consequence of a different toxicity or adverse events of the treatment itself. Identifying this effect provides relevant information particularly in those cases, as the one presented in the case study discussed in Section 5, where the improvement in the survival (or in the primary end point of the trial) due to a new treatment is moderate. In this case, to expose patients to a new treatment may be ethically and economically appropriate if this does not have a detrimental effect on their HRQOL. On the other hand, adopting a new treatment can be inappropriate for those patients who, being in a poorer HRQOL state, may suffer side effects as a result of the new treatment beyond what is considered to be

ethical, given the expected improvement in the survival. Some further actions should have been considered for the latter kind of patients who have been medicated with the new treatment, such as explaining to them the potential consequences of the therapy, planning appropriate supporting actions when in therapy and so on.

This type of analysis is not possible by using conventional regression models for the conditional expectation but it can be obtained using quantile regression models. Quantile regression was introduced in the econometrics literature by Koenker and Bassett (1978) and thoroughly described by Koenker (2005). Since then, quantile regression has been applied in many and different areas of research. Recently it has been successfully employed in medical applications (Carey et al., 2004; Austin et al., 2005; Wei et al., 2006; Geraci and Bottai, 2007; Bottai et al., 2010; Li et al., 2010). Only few applications of quantile regression to modelling HRQOL currently exist in literature (Nicholson et al., 2006; Pourhoseingholi et al., 2008; Broccoli et al., 2005), but the studies do not take into account the hierarchical structure of the data. Recently, Geraci and Bottai (2007, 2014) propose a conditional two-levels quantile regression model (LQMM) that assumes an Asymmetric Laplace Distribution (ALD) for modelling the conditional likelihood given the random effects. Inference for the model parameters is performed by using a bootstrap approach based on resampling the sample data. Estimation and inference is facilitated by the `lqmm` package in `R`.

There are, however, alternatives to quantile regression, such as M-quantile regression (Breckling and Chambers, 1988), which is a quantile-like generalization of regression based on influence functions (M-regression), and expectile regression (Newey and Powell, 1987), a quantile-like generalization of mean regression. Tzavidis et al. (2015) propose the extension of M-quantile regression to two-levels M-quantile random effects regression for multilevel-type data.

This paper is, to the best of our knowledge, the first attempt to apply a quantile-like random effects model to complex data from longitudinal and multi-centre studies (three-levels models). More specifically, in this paper we proposes the extension of two-levels M-quantile regression to three-levels M-quantile regression models.

M-quantile regression (Breckling and Chambers, 1988) integrates the concepts of quantile regression and expectile regression within a framework defined by a 'quantile-like' generalization of regression based on influence functions (M-regression). The M-quantile of order q for the conditional density of y given the set of covariates $\mathbf{X}$, $f(y|\mathbf{X})$, is defined as the solution $MQ_y(q|\mathbf{X};\psi)$ of the estimating equation $\int \psi_q\Big\{y - MQ_y(q|\mathbf{X};\psi)\Big\}f(y|\mathbf{X})dy = 0$, where ψ_q denotes an asymmetric influence function, which is the derivative of an asymmetric loss function ρ_q. A linear M-quantile regression model for y_i given $\mathbf{x}_i$ is one where we assume that

$$MQ_{y_i}(q|\mathbf{x}_i;\psi) = \mathbf{x}_i^T\boldsymbol{\beta}_q,$$

and estimates of $\boldsymbol{\beta}_q$ are obtained by minimising

$$\sum_{i=1}^{n}\rho_q\{r_{iq}\}, \tag{1}$$

where $r_{iq} = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}_q}{\sigma}$, σ is a scale parameter and $\rho_q\{r_{iq}\} = 2\rho\{r_{iq}\}\Big[qI(r_{iq} > 0) + (1 - q)I(r_{iq} \leqslant 0)\Big]$. Different set of regression parameters can be defined for each value of q. In particular, by varying the specifications of the asymmetric loss function ρ we obtain the expectile, M-quantile and quantile regression models as special cases. When ρ_q is the square loss function, we obtain the linear expectile regression model if $q \neq 0.5$ (Newey and Powell, 1987) and the standard linear regression model if $q = 0.5$. When ρ is the loss function described by Koenker and Bassett (1978) we obtain the linear quantile regression. Throughout this paper we will take the linear M-quantile regression model to be defined by using as ρ the Huber loss function (Huber, 1981). For more details on M-quantile regression models see Breckling and Chambers (1988).

Quantiles have a more intuitive interpretation than M-quantiles even if they target essentially the same part of the distribution of interest. In this paper we use M-estimation because it offers some advantages: (i) it easily allows for robust estimation of both fixed and random effects; (ii) it can trade robustness for efficiency in inference by selecting the tuning constant of the influence function; (iii) it can offer computational stability because it can use a wide range of continuous influence functions instead of the absolute value used in the quantile regression.

Tzavidis et al. (2015) extended M-quantile regression to include random effects to account for a two-level hierarchical structure in the data and used maximum likelihood to estimate the parameters of the model. In this paper we extend this approach to a three-level random effects model, which is appropriate for the analysis of the data we consider in Section 2 and we propose a maximum likelihood approach to estimate the model parameters. We applied the proposed methodology to a prospective, randomized, multi-centre clinical trial. The paper is structured as follows. The data are introduced in Section 2. Section 3 presents the proposed methodology. In Section 4 we evaluate the proposed regression models using model-based simulation studies, under a range of different data generating mechanisms. In Section 5 we present the results from the application of three-levels, two-levels M-quantile random effects regression models and quantile random effects regression (Geraci and Bottai, 2014) to the HRQOL data. The results are discussed and concluding remarks are presented in Section 6.

2 The data: HRQOL of advanced melanoma patients in a multi-centre clinical trial

The study considered in this paper is a prospective, randomized, multi-centre phase III clinical trial that aimed at comparing the efficacy of two treatments, Cisplatin 75 mg/m^2 and DTIC 800 mg/m^2 (CT) versus the same regimen plus IL-2 and IFN-2b (bio-CT), in advanced melanoma patients, who had not been previously treated with systemic chemotherapy. Both treatments were administered for six cycles or until disease progression. The primary objective of the trial was overall survival, while HRQOL evaluation was planned as a secondary objective. Further details on the clinical analysis and the

Table 1: Summary statistics of PSDS stratified by treatment.

		Order of sample quantiles						
treatment	N. Obs	0.10	0.25	0.50	0.75	0.90	mean	std. dev.
CT	264	68.70	78.26	87.88	93.94	97.10	84.73	11.82
bio-CT	244	64.26	75.76	86.66	92.75	97.10	83.17	12.72

Table 2: Summary statistics of PSDS at different temporal occasions.

		Order of sample quantiles						
temporal occasions	N. Obs	0.10	0.25	0.50	0.75	0.90	mean	std. dev.
0	137	69.70	82.61	90.91	96.97	98.55	87.80	11.25
1	121	65.08	76.81	86.36	92.42	95.65	83.20	12.02
2	98	63.64	74.89	86.04	92.32	97.10	83.00	12.36
3	68	61.84	75.76	84.85	89.86	94.36	81.07	13.07

HRQOL analysis are reported elsewhere (Ridolfi et al., 2002; Chiarion-Sileni et al., 2003).

The HRQOL status was assessed by a self-reported questionnaire, the Rotterdam Symptom Checklist (RSCL). In this paper, we focus on the physical symptom distress scale (PSDS) score as the primary HRQOL outcome. The questionnaire was administered to all patients for completion prior to the first cycle of chemotherapy (baseline assessment), and subsequently just before each successive cycle of chemotherapy. The HRQOL evaluation was not planned after disease progression or during the follow-up period. Since a large part of the sample experienced disease progression starting from the fourth assessment onwards, we limited our analysis to the data collected in the first four occasions, since the sample became too small after this assessment.

The data were collected between March 1997 and December 1999. The trial enrolled 178 patients from 23 different clinical centres, half of them were randomized to receive CT and the remaining 89 to receive bio-CT. The median time to progression was 3.6 months for bio-CT and 3 months for CT, showing a moderate effect of the treatment under evaluation.

We considered 137 patients for the HRQOL analysis discarding 16 patients who did not have a baseline measurement, 22 patients who never completed any form and three who died before the start of treatment. In total there were 508 measurements of PSDS, the average being 83.98 (sd=12.27). Summary statistics of PSDS stratified by treatment are shown in Table 1. A difference between the two groups is evident in the left tail of the PSDS distribution whereas the discrepancy is negligible at the centre as well as the right tail of it.

Table 2 shows key sample quantiles of PSDS stratified by the temporal occasion at which patients were evaluated. From Figure 1 it clearly appears that the HRQOL, as measured by PSDS, tended to decline, roughly linearly, at all quantiles as the study progresses. As mentioned above, the number of patients in the study declined as the study continued.

Figure 2 shows the centre-specific distribution of PSDS scores at baseline. The distribution appears to vary from centre to centre, suggesting that a certain degree of variability in PSDS measures can be attributed to the centres which participated in the trial.

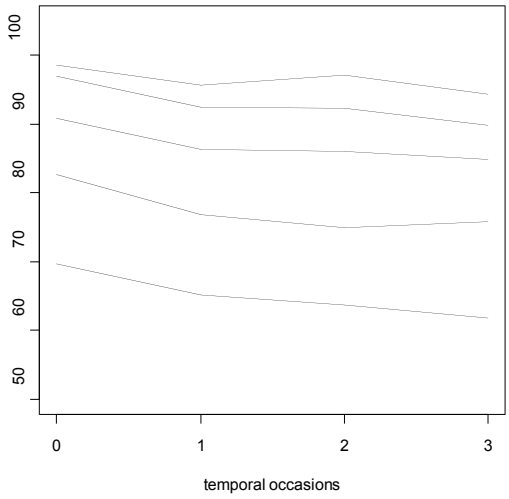

Figure 1: PSDS quantiles versus temporal occasions. The quantiles considered are 0.10, 0.25, 0.50, 0.75 and 0.90.

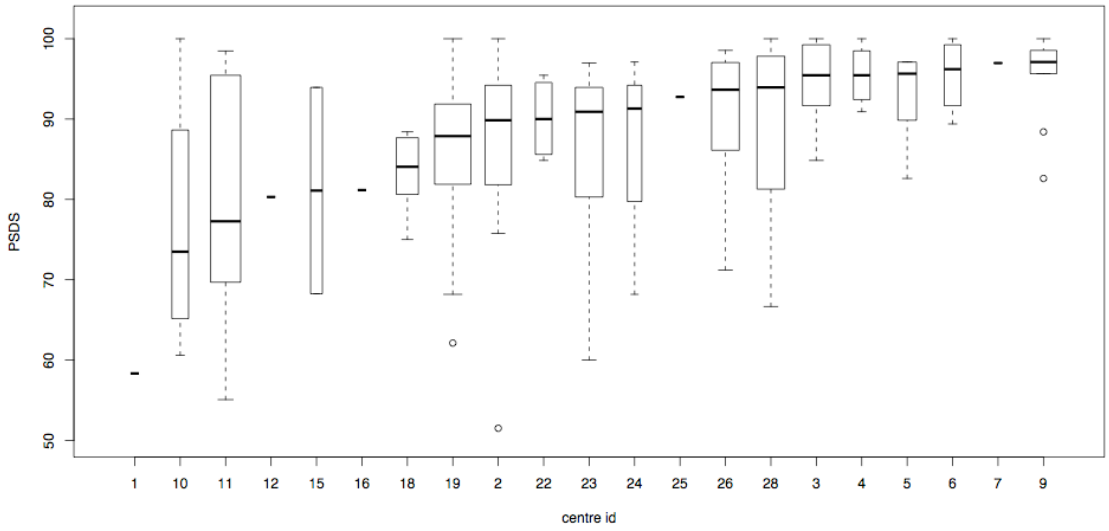

Figure 2: Distribution of PSDS scores at baseline. The width of each box is proportional to the number of patients in each centre showing the different sample sizes in each of them.

3 Three-level M-quantile random effects regression

In what follows we present a three-level M-quantile random effects regression. In order to bring continuity, we first briefly review random effects models. We consider data with a three-level hierarchical structure because in the multi-centre clinical trial data set, occasions (repeated measures denoted by k - level 1) are nested within patients (denoted by j - level 2), and patients are nested within centres (denoted by i - level 3). A conventional three-level random effects model is described by

$$y_{ijk} = \mathbf{x}_{ijk}^T \boldsymbol{\beta} + u_i + \gamma_{ij} + \varepsilon_{ijk},\ i = 1, \ldots, m,\ j = 1, \ldots, n_i,\ k = 1, \ldots, t_{ij}, \qquad (2)$$

where y_{ijk} denote the value of the variable of interest y, $\mathbf{x}_{ijk}$ is a p vector of auxiliary variables associated with measure k of unit j in group i and $\sum_{i=1}^{m} n_i = n$ and $\sum_{i=1}^{m}\sum_{j=1}^{n_i} t_{ij} = N$. We assume that $\mathbf{x}_{ijk}$ contains 1 as its first component. Here $\boldsymbol{\beta}$ is a p vector of fixed effect parameters, ε_{ijk} is the occasion random effect, γ_{ij} is a patient random effect and u_i denotes a m vector of centre specific random effects. Model (2) conventionally assumes that $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u)$, with $\boldsymbol{\Sigma}_u = \sigma_u^2 \mathbf{I}_m$, $\boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\gamma)$, with $\boldsymbol{\Sigma}_\gamma = \sigma_\gamma^2 \mathbf{I}_n$, $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$ and mutually independent. Here and throughout the paper $\mathbf{I}_g$ is an identity matrix of size g. In matrix form model (2) can be re-written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\mathbf{u} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{3}$$

where $\mathbf{Z}$ and $\mathbf{D}$ are an auxiliary contextual matrices. In simplest case, we assume $\mathbf{Z}$ and $\mathbf{D}$ as incidence matrices which specify the random intercepts specification of model (2). Maximum likelihood (ML) estimation based on the marginal distribution of y is widely used for estimating the unknown parameters of model (2) (Harville, 1977). In particular, estimates of $\boldsymbol{\beta}$, σ_u^2, σ_γ^2 and σ_ε^2 are obtained by first differentiating the log-likelihood with respect to these parameters and then solving the estimating equations defined by setting these derivatives equal to zero (Goldstein, 2003). It is easy to see that in this log-likelihood a squared loss function of the residuals is assumed. Estimates of the random effects are then obtained by using the maximum likelihood estimates of the fixed effects and the variance components.

Data may contain outliers that invalidate the Gaussian assumptions. In such a case, the estimated model parameters under (3) will be biased and inefficient (Richardson and Welsh, 1995). A number of papers (Huggins, 1993; Richardson and Welsh, 1995; Richardson, 1997; Huggins and Loesch, 1998) proposed robust estimation of the random effects model, which offers protection against departures from normality. This is achieved by using an alternative loss function in the log-likelihood that grows along with the regression residuals at a slower rate than the squared loss function. Alternatively, Richardson and Welsh (1995) proposed a robust version of the estimating equations of the log-likelihood function, but there is no associated likelihood function unlike in the case of large tuning constant in the influence function. For details see Richardson and Welsh (1995) and Richardson (1997).

For describing the relationship between y and a set of covariates $\mathbf{X}$ at other parts of a conditional distribution we extend the two level M-quantile random effects regression model (Tzavidis et al., 2015) to a three-level M-quantile random effects regression model. In particular, we propose using asymmetric loss functions for this purpose when the data are hierarchically structured. Since the estimating equations obtained from the modified marginal log-likelihood function (Richardson and Welsh, 1995; Richardson, 1997) are susceptible to multiple solutions, we begin from the robust maximum likelihood proposal II in Richardson and Welsh (1995). According to Sinha and Rao (2009), we observe that one can extend the idea of asymmetric weighting of residuals to hierarchical data by defining

the following modified estimating equations

$$\mathbf{X}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} = \mathbf{0} \tag{4}$$

$$\begin{aligned}
\frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2}tr\left[\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\right] &= 0 \\
\frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2}tr\left[\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\right] &= 0 \\
\frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2}tr\left[\mathbf{V}_q^{-1}\right] &= 0,
\end{aligned} \tag{5}$$

where $\mathbf{r}_q = \mathbf{U}_q^{-1/2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_q)$ is a vector of scaled residuals with components r_{ijkq}, $\mathbf{U}_q$ is a diagonal matrix with diagonal elements u_{ijq} equal to the diagonal elements of the covariance matrix $\mathbf{V}_q$ and $\psi_q(r) = 2\psi(r)\{qI(r > 0) + (1-q)I(r \leq 0)\}$ is a bounded influence function obtained as the derivative of a loss function ρ_q. Here $\mathbf{V}_q = \mathbf{D}\boldsymbol{\Sigma}_{u_q}\mathbf{D}^T + \mathbf{Z}\boldsymbol{\Sigma}_{\gamma_q}\mathbf{Z}^T + \boldsymbol{\Sigma}_{\epsilon_q}$, $\boldsymbol{\Sigma}_{u_q} = \sigma^2_{u_q}\mathbf{I}_m$, $\boldsymbol{\Sigma}_{\gamma_q} = \sigma^2_{\gamma_q}\mathbf{I}_n$, $\boldsymbol{\Sigma}_{\epsilon_q} = \sigma^2_{\epsilon_q}\mathbf{I}_N$, σ_{u_q}, σ_{γ_q} and σ_{ε_q} are the quantile-specific variance parameters, and $\boldsymbol{\beta}_q$ is the $p\times 1$ vector of M-quantile regression coefficients. Finally, the component $K_2 = E[\psi(\boldsymbol{\varepsilon})\psi(\boldsymbol{\varepsilon})^T]$ with $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{I}_N)$.

With (5) we extend the idea of weighting positive residuals by q and negative residuals by $(1-q)$, where $0 < q < 1$ is the quantile order, used in fitting the single-level M-quantile regression (Huber, 1981), to M-quantile regression for hierarchically structured data.

We note that the ML equations of model (3) and their robust version are a special case of the estimating equations in (5) for a specific choice of ρ_q and q. For example, when $q = 0.5$ and ρ_q is the squared loss function we obtain ML estimating equations whereas when $q = 0.5$ and we use a loss function other than squared loss, for example the Huber loss function, we obtain a robust version of the ML proposal II estimating equations by Richardson and Welsh (1995). For q values other than 0.5 and for different choices of ρ_q, the solutions of (5) will provide estimates of fixed effects and variance parameters, which can then be used for obtaining the M-quantile random effects regression (MQRE) fits. In particular, using a squared loss function and its derivative (influence function) in (5) at $q \neq 0.5$ results in an expectile random effects fit.

As discussed in Jones (1994), under specific distributions the relationship between quantiles and M-quantiles is known and both quantiles and M-quantiles model the same part of the distribution of interest. In this case estimates of the fixed effects $\boldsymbol{\beta}_q$ can be practically interpreted, for example, as the effect of a one unit increase in x on the lower quartile, middle or upper quartile of the distribution of y. The variance parameters show the between and within group dispersion around the M-quantile being estimated.

Estimating equations can be solved iteratively to obtain estimators of $\boldsymbol{\beta}_q$, σ_{u_q}, σ_{γ_q} and σ_{ε_q}. Here we adopt Newton-Raphson algorithm for solving (4) and the fixed-point iterative method (Anderson, 1973) for solving estimating equations (5). The steps of the estimation algorithm are outlined in Appendix A. Robust estimates of the random effects can be obtained by solving a modified version of the estimating equations proposed by Fellner (1986) at each value of q. See Appendix A and Tzavidis et al. (2015) for details.

A function that fits the three-levels M-quantile (expectile) random effects regression has been in written in R. Some asymptotic properties of the estimators and their variance parameters for the MQRE models of order q are discussed in Tzavidis et al. (2015).

4 A simulation experiment

In order to compare the performance of the estimators from Section 3, a Monte-Carlo simulation was conducted. In this simulation we evaluate the performance of the MQRE (three-level) at two quantiles, $q = 0.25$ and $q = 0.5$. The aim here is two-fold. For one thing, we investigate the ability of the MQRE to account for the dependence structure of hierarchical data beyond two-levels and for another thing, we assess the asymptotic approximations of the standard errors of the regression parameters and the variance parameters (level 2 and level 3). For both goals, a nested error regression model

$$y_{ijk} = 100 + 2\boldsymbol{x}_{ijk} + u_i + \gamma_{ij} + \varepsilon_{ijk} \;\; i = 1, ..., 40, \;\; j = 1, ..., 200, \;\; k = 1, ..., n_j \tag{6}$$

is used as a core model for the generation of the dependent variable. Similar to the HRQOL data, the sample sizes n_j vary between 1 and 5 leading to a total sample size of $n = 505$. The auxiliary variables are uniformly distributed in $[0, 20]$. The covariates and the sample sizes are held fixed throughout the simulation study. The level 1, level 2 and level 3 error terms ε_{ijk}, γ_{ij} and u_i are independently generated depending on three different scenarios:

- Normal distribution ($[N, N, N]$): $u_i \sim N(0, 1)$, $\gamma_{ij} \sim N(0, 4)$ and $\varepsilon_{ijk} \sim N(0, 16)$,
- t-distribution ($[N, N, T]$): $u_i \sim N(0, 1)$, $\gamma_{ij} \sim N(0, 2)$ and $\varepsilon_{ijk} \sim t(df = 3)$,
- Outliers in level 1 and 2 ($[N, \gamma, \varepsilon]$): $u_i \sim N(0, 1)$, $\gamma_{ij} \sim 0.9N(0, 4) + 0.1N(0, 20)$ and $\varepsilon_{ijk} \sim 0.9N(0, 16) + 0.1N(0, 150)$.

Each scenario is evaluated by $R = 500$ independent Monte-Carlo replications. The assumptions of the underlying three-level random effects model (2) are valid under scenario $[N, N, N]$. In setting $[N, N, T]$ we have some departures from normality by a t-distribution, whereas scenario $[N, \gamma, \varepsilon]$ displays a setting under a mixture distribution (outlier contamination) in level 1 and 2. The tuning constant c of the Huber influence function is set to 1.345 for the MQRE.

Starting with the first aim of the simulation section, we assess the performance of the MQRE with 3 levels (MQRE), the MQRE with 2 levels (MQRE-2L) regression model (Tzavidis et al., 2015) and the linear M-quantile (MQ) regression model (Breckling and Chambers, 1988). We expect that the MQRE performs on a higher level compared to the MQRE-2L and MQ when clustering occurs in the data for quantiles $q = 0.25$ and $q = 0.5$. At $q = 0.5$, we compare the MQRE also with the two-level linear random effects model (LRE-2L) and the three-level linear random effects model (LRE). We suppose more efficient results of the LRE compared to the MQRE in the setting under normality.

Table 3: Values of bias (ARB), efficiency (EFF), and average of point estimates over simulations of fixed effects under the three data generating scenarios and the alternative regression fits: MQRE, MQRE-2L, MQ, LRE, LRE-2L at q = 0.25 and q=0.5.

	$\hat{\beta}_0$			$\hat{\beta}_1$			$\hat{\beta}_0$			$\hat{\beta}_1$		
	ARB	EFF	$\bar{\beta}_0$	ARB	EFF	$\bar{\beta}_1$	ARB	EFF	$\bar{\beta}_0$	ARB	EFF	$\bar{\beta}_1$
	$q = 0.25$						$q = 0.5$					
	Scenario 1 - $[N, N, N]$						Scenario 1 - $[N, N, N]$					
MQRE	0.558	0.979	97.845	0.057	0.956	2.001	-0.011	0.968	99.989	-0.004	0.956	2.000
MQRE-2L	0.570	0.996	97.857	0.010	0.992	2.000	-0.000	0.990	100.000	-0.043	0.986	1.999
MQ	0.571	1.000	97.857	0.017	1.000	2.000	0.003	1.000	100.003	-0.051	1.000	1.999
LRE	—	—	—	—	—	—	-0.002	0.941	99.998	-0.018	0.925	2.000
LRE-2L	—	—	—	—	—	—	0.007	0.971	100.007	-0.054	0.970	1.999
	Scenario 2 - $[N, N, T]$						Scenario 2 - $[N, N, T]$					
MQRE	-0.291	0.784	98.946	0.024	0.630	2.000	-0.013	0.760	99.987	0.034	0.572	2.001
MQRE-2L	-0.290	0.902	98.947	0.011	0.853	2.000	-0.010	0.888	99.990	0.005	0.834	2.000
MQ	-0.289	1.000	98.949	-0.000	1.000	2.000	-0.005	1.000	99.995	-0.024	1.000	2.000
LRE	—	—	—	—	—	—	-0.008	0.835	99.992	0.019	0.717	2.000
LRE-2L	—	—	—	—	—	—	-0.006	0.957	99.994	0.001	0.932	2.000
	Scenario 3 - $[N, \gamma, \varepsilon]$						Scenario 3 - $[N, \gamma, \varepsilon]$					
MQRE	0.258	0.964	97.553	0.027	0.962	2.001	0.010	0.961	100.010	-0.014	0.953	2.000
MQRE-2L	0.263	0.989	97.558	0.004	0.988	2.000	0.015	0.983	100.015	-0.044	0.992	1.999
MQ	0.266	1.000	97.561	0.008	1.000	2.000	0.011	1.000	100.011	-0.023	1.000	2.000
LRE	—	—	—	—	—	—	0.010	1.146	100.011	-0.014	1.188	2.000
LRE-2L	—	—	—	—	—	—	0.014	1.153	100.014	-0.036	1.218	1.999

In contrast, the MQRE should perform on a higher level compared to the LRE in the scenarios with contamination. The focus of the first part of the simulation study is set on comparing fixed effects of the different methods. The results for the variance parameters are available from the authors upon request. For each regression parameter performance is evaluated using the following quality measures:

(a) Average Relative Bias (ARB) defined as

$$ARB(\hat{\theta}) = R^{-1} \sum_{r=1}^{R} \frac{\hat{\theta}^{(r)} - \theta}{\theta} \times 100,$$

where $\hat{\theta}^{(r)}$ is the estimated parameter at quantile q for the rth replication and θ is the corresponding 'true' value of this parameter.

(b) Relative Efficiencies (EFF) defined as

$$EFF(\hat{\theta}) = \frac{S^2_{model}(\hat{\theta})}{S^2_{\mathsf{MQ}c=1.345}(\hat{\theta})}$$

where $S^2(\hat{\theta}) = R^{-1} \sum_{r=1}^{R} (\hat{\theta}^{(r)} - \bar{\theta})^2$ and $\bar{\theta} = R^{-1} \sum_{r=1}^{R} \hat{\theta}^{(r)}$.

Table 3 shows the simulation results for the different methods of the fixed effects under various scenarios for quantiles $q = 0.25$ and $q = 0.5$. The tables indicates that the estimators from LRE are more efficient than the corresponding estimator from the MQRE under the scenario $[N, N, N]$ for quantile $q = 0.5$. Under this scenario, there is no reason to use outlier-robust methods and therefore, this leads to a higher variability of the MQRE regression estimators. It can also be observed that the MQRE is more efficient

Table 4: Empirical standard errors and estimated standard errors of $\hat{\beta}_q$, $\hat{\sigma}_u^2$ and $\hat{\sigma}_\gamma^2$ for $q = (0.25, 0.5)$ using MQRE for the three-level model.

	Empir. s.e.	Estim. s.e.	Empir. s.e.	Estim. s.e.	Empir. s.e.	Estim. s.e.	Empir. s.e.	Estim. s.e.
	$\hat{\beta}_0$		$\hat{\beta}_1$		$\hat{\sigma}_u^2$		$\hat{\sigma}_\gamma^2$	
MQRE	$q = 0.25$							
Scen. 1 - $[N, N, N]$	0.4675	0.4644	0.0365	0.0362	0.4070	0.4894	0.9223	0.9473
Scen. 2 - $[N, N, T]$	0.2414	0.2528	0.0149	0.0149	0.2340	0.2586	0.2791	0.3403
Scen. 3 - $[N, \gamma, \varepsilon]$	0.5729	0.5432	0.0470	0.0433	0.4741	0.5996	1.2464	1.3284
MQRE	$q = 0.5$							
Scen. 1 - $[N, N, N]$	0.4481	0.4435	0.0349	0.0339	0.6527	0.7178	1.2731	1.2512
Scen. 2 - $[N, N, T]$	0.2287	0.2361	0.0135	0.0131	0.3474	0.2969	0.3715	0.3261
Scen. 3 - $[N, \gamma, \varepsilon]$	0.5177	0.4960	0.0411	0.0388	0.7541	0.8545	1.5944	1.6716

than the corresponding MQRE-2L or MQ. This is natural, because the MQRE correctly models the three-level structure present in the synthetic population. Coming now to quantile $q = 0.25$, the outstanding performance of the MQRE is especially demonstrated in settings with clear departures from normality. For instance, the regression parameters of the MQRE have a smaller standard error than the corresponding estimators from two-level models (MQRE-2L or LRE-2L) or single level models (MQ). Furthermore, Table 3 indicates in the settings under contamination ($[N, N, T]$ and $[N, \gamma, \varepsilon]$) that the MQRE is more efficient than the MQRE-2L. These phenomena show that using the MQRE with 3 levels protects against outlying values and it accounts for the specific dependence structure of hierarchical data. Coming now to the average relative bias (ARB) of the regression parameters of the different methods, Table 3 reveals that all methods are almost unbiased in all scenarios. For instance, we observe an ARB of less than 0.1% for the intercept and slope for all estimation methods for $q = 0.5$.

Having evaluated the efficiency and bias of MQRE, the second aim of this Section is to assess the asymptotic approximations of the standard errors of the fixed effects and the variance parameters. Thus, we compare the estimated and empirical standard errors under the three scenarios. For each scenario and for each estimator $\hat{\theta}$, at $q = 0.25$ and $q = 0.5$, Table 4 reports the estimated standard errors of the fixed effects $\boldsymbol{\beta}$ and variance parameters, $\hat{\sigma}_u^2$ and $\hat{\sigma}_\gamma^2$, and the Monte-Carlo standard error $S(\hat{\theta}) = \sqrt{R^{-1} \sum_{r=1}^{R} (\hat{\theta}^{(r)} - \bar{\theta})^2}$. The table reveals that for all scenarios the asymptotic standard error of the fixed effects and the variance parameters at $q = 0.25$ and $q = 0.5$ provides a good approximation to the true variances. However, in some cases we observe slightly under- or overestimation, especially for the intercept and the variance parameter $\hat{\sigma}_u^2$. We expect that the results can potentially be improved as the number of observations within groups, the number of groups and the number of Monte-Carlo replications are raised.

Overall, the MQRE with three-levels is a good compromise between efficiency under normality and robust properties under contamination. Furthermore, the MQRE performs on a higher level than the MQRE-2L when three-level clustering is present in the data.

5 M-quantile modelling of HRQOL

In this section we apply the methodology described in Section 3 for modelling the distribution of HRQOL in advanced melanoma patients using the data of the randomized multi-centre clinical trial presented in Section 2. The aim of the trial was to compare the efficacy of CT versus bio-CT as a primary end-point. Here we considered the impact of the two treatments on the HRQOL. In particular we modeled the change of the HRQOL score at each time point from the baseline as a function of the treatment. In addition to the treatment effect, the model includes a linear trend since a reduction in the quality of life is expected as the exposure to the treatment increases. To account for the hierarchical nature of the data we also included a random effect to capture the between centre variability and a random effect to account for unobserved individual heterogeneity, i.e. the variability due to repeated measurements on the same patient.

5.1 Preliminary data analyses

Before presenting the results from the study we present some preliminary data analysis. We applied the following three-level model for the analysis:

$$y_{ijk} = \beta_0 + \beta_1 bio\text{-}CT_{ijk} + \beta_2 temporal\ occasion_{ijk} + centre\ id_i + npz_{ij} + \varepsilon_{ijk},$$

where, as said before, y_{ijk} is the change of the HRQOL score at each time point from the baseline. Figure 3 shows normal probability plots of level 1 - plot (a) -, level 2 - plot (b) - and level 3 - plot (c) - residuals obtained by fitting a three-level mixed model to the data. We denote by level 1 the observation at a given temporal occasion, by level 2 the patient level (*npz*) and by level 3 the *centre*. The normal probability plots indicate that the Gaussian assumptions of the mixed model are not met. This is confirmed by a Shapiro-Wilk normality test, which rejects the null hypothesis that the residuals follow a normal distribution (p-values: level 1 = 2.42e-06, level 2 = 0.05771, level 3 = 0.02672).

Figure 4 shows two plots of standardized mixed model residuals. The histogram, plot (a), depicts a skewed distribution of the residuals. This is confirmed by plot (b) which shows the distribution of standardized mixed model residuals by centre: some centres have many positive residuals, whereas others have many negative residuals. From this second plot we can observe some high residuals r in absolute value ($|r| > 2$). This indicates the presence of influential observations in the data. Figure 5 shows the Cook's Distance by centre, plot (a), and by patient, plot (b). These graphs also suggest the presence of influential observations in the data. Our preliminary analysis hence indicates that the estimates of model parameters can be potentially driven by the influential observations and hence robust estimation may improve inference under the random effects model.

5.2 Results

Table 5 shows the estimated three-levels MQRE regression coefficients for five quantiles: 0.1, 0.25, 0.5, 0.75 and 0.90. The tuning constant c was fixed at 1.345. This value gives

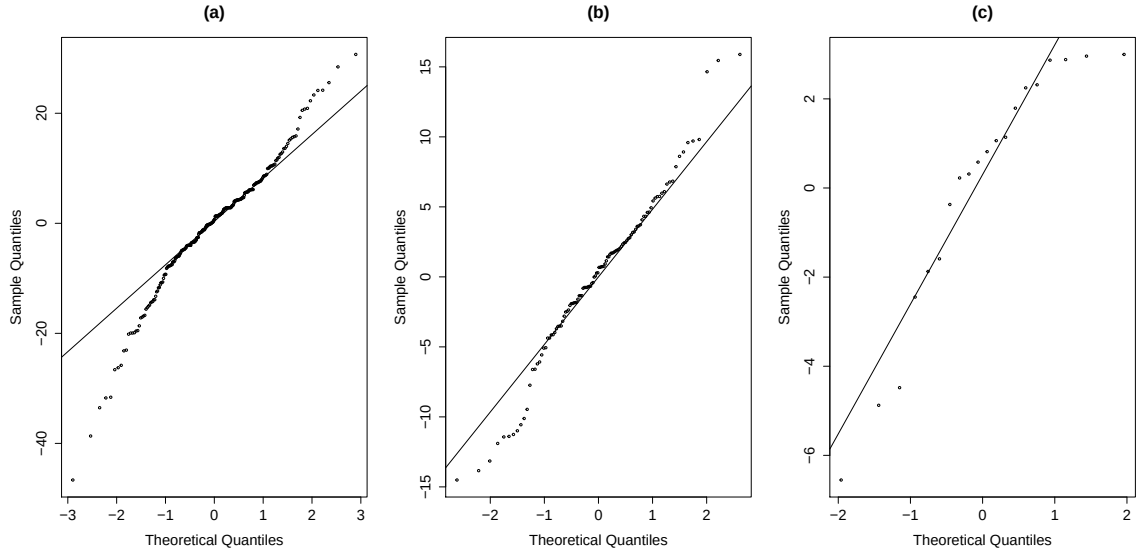

Figure 3: Normal probability plots of level 1 (a) level 2 (b) and level 3 residuals (c) derived by fitting a three-levels linear mixed model to the survey data.

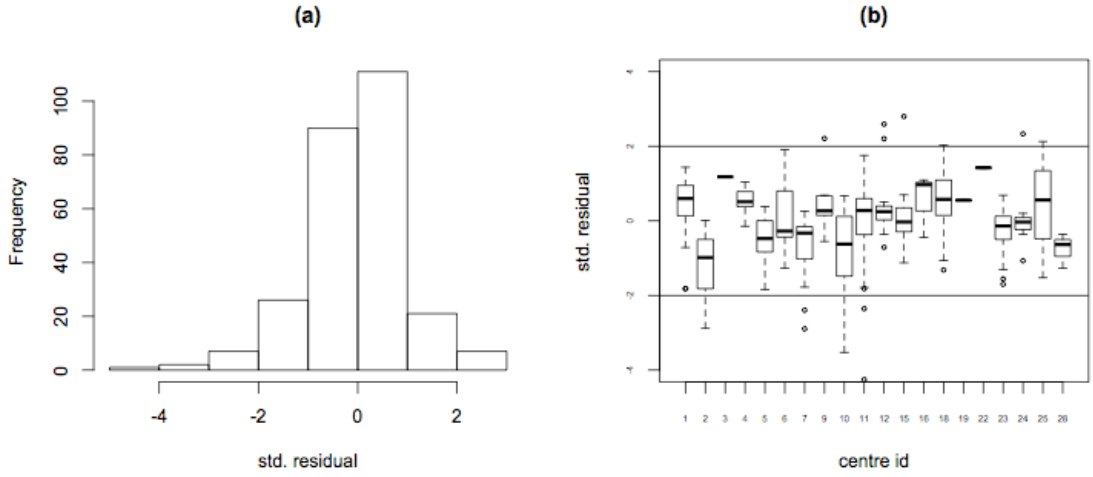

Figure 4: A histogram of standardized mixed model residuals (a) and the distribution of standardized mixed model residuals by centre (b).

reasonably high efficiency under normality and protects against outliers (Huber, 1981). The estimates of the bio-CT coefficient measure the effect of bio-CT, as opposed to CT, on the PSDS at a given quantile.

The intercept of the model associated with each conditional quantile represents the value of PSDS at the first occasion, i.e. the effect of the first cycle of chemotherapy on each patient after controlling for the centre and patient heterogeneity.

The plots in Figure 6 show the estimated effect of each explanatory variable we included in the model by quantiles. Estimates far from the centre of the distribution usually cannot be evaluated with high precision. To display the sampling variation, a confidence band across the quantiles was constructed by estimating the point-wise 95% confidence interval for the regression coefficients associated with the selected quantiles. Grey-shaded areas around the line represent confidence bands. It appears that variation differs among quantiles, sometimes substantially, and generally increases as the quantile

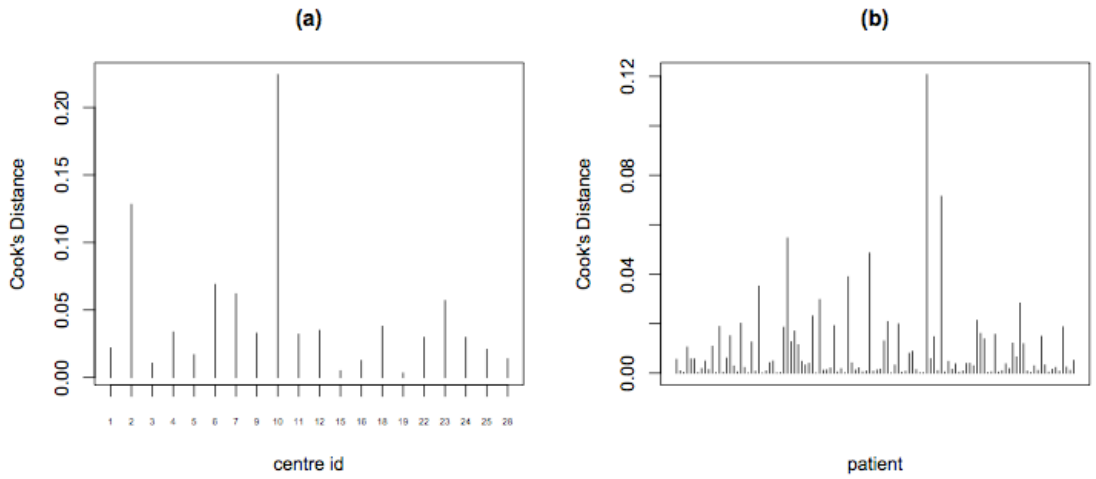

Figure 5: Cook's Distance by centre (a) and by patient (b).

Table 5: Parameter estimates and corresponding standard error estimates in parentheses for the data. MQRE three-level.

	q=0.10	q=0.25	q=0.50	q=0.75	q=0.90
intercept	-7.69 (2.56)	-4.51 (1.73)	-1.91 (1.30)	0.67 (1.58)	2.63 (2.51)
bio-CT	-4.37 (2.62)	-3.52 (1.58)	-2.99 (1.06)	-2.64 (1.06)	-2.10 (1.98)
temporal occasions	-3.33 (1.48)	-2.25 (0.87)	-0.97 (0.51)	-0.15 (0.56)	0.48 (0.88)
$\sigma^2_{centre_q}$	3.24 (4.08)	6.58 (7.51)	11.97 (7.66)	5.19 (2.59)	0.08 (1.99)
$\sigma^2_{npz_q}$	17.10 (9.28)	35.32 (15.14)	33.03 (7.86)	27.81 (12.05)	20.32 (11.70)
$\sigma^2_{\varepsilon_q}$	32.28 (13.52)	36.25 (10.00)	22.99 (4.95)	17.39 (6.39)	14.59 (6.19)

Table 6: Parameter estimates and corresponding standard error estimates in parentheses for the data. MQRE two-level.

	q=0.10	q=0.25	q=0.50	q=0.75	q=0.90
intercept	-8.06 (3.18)	-4.71 (1.94)	-2.20 (1.29)	0.27 (1.51)	2.61 (2.52)
bio-CT	-4.00 (3.04)	-3.34 (2.07)	-2.88 (1.61)	-2.46 (1.72)	-2.10 (2.56)
temporal occasions	-3.43 (1.58)	-2.32 (0.93)	-1.06 (0.54)	-0.22 (0.56)	0.47 (0.86)
$\sigma^2_{npz_q}$	19.10 (9.93)	42.07 (13.72)	45.10 (8.83)	31.99 (12.50)	20.53 (12.33)
$\sigma^2_{\varepsilon_q}$	32.11 (14.78)	36.73 (9.64)	24.10 (4.88)	16.36 (5.40)	14.65 (5.46)

Table 7: Parameter estimates and corresponding standard error estimates in parentheses for the data. LQMM (two-level) - random intercepts.

	q=0.10	q=0.25	q=0.50	q=0.75	q=0.90
intercept	-10.22 (2.53)	-6.38 (2.13)	-0.65 (2.14)	0.72 (1.68)	2.38 (2.65)
bio-CT	-1.92 (4.07)	-4.45 (2.43)	-2.83 (2.88)	-1.52 (1.99)	-1.76 (3.05)
temporal occasions	-1.90 (1.13)	-1.29 (0.83)	-0.86 (0.61)	-0.72 (0.60)	-0.64 (0.67)
$\sigma^2_{npz_q}$	77.21 (—)	61.33 (—)	50.96 (—)	52.52 (—)	48.46 (—)

order approaches 0 or 1. On the edges of the probability range, such an increase can be quite large as shown in Figure 6.

Table 5 shows that the centre effect ($\sigma^2_{centre_q}$) is low at the tail of the distribution and it is high in the centre of the distribution. However this effect is not significant for all the locations. As it has also mentioned in the introduction in randomized clinical trial studies, patients are recruited at multiple centers to accrue large enough samples within an acceptable period. This rises an issue concerning the heterogeneity induced by potentially

Table 8: Parameter estimates and corresponding standard error estimates in parentheses for the data. LQMM (two-level) - random slopes.

	q=0.10	q=0.25	q=0.50	q=0.75	q=0.90
intercept	-7.27 (2.70)	-5.30 (2.24)	-1.34 (2.21)	1.45 (1.66)	2.81 (2.58)
bio-CT	-3.88 (4.10)	-1.82 (2.50)	-4.10 (2.68)	-0.89 (1.87)	-0.30 (2.73)
temporal occasions	-2.49 (1.09)	-3.02 (0.87)	-1.50 (0.64)	-1.45 (0.57)	-1.42 (0.80)
$\sigma^2_{npz_q}$	59.94 (—)	55.20 (—)	31.78 (—)	29.46 (—)	49.78 (—)
σ^2_{time}	10.40 (—)	10.73 (—)	7.68 (—)	7.04 (—)	5.50 (—)

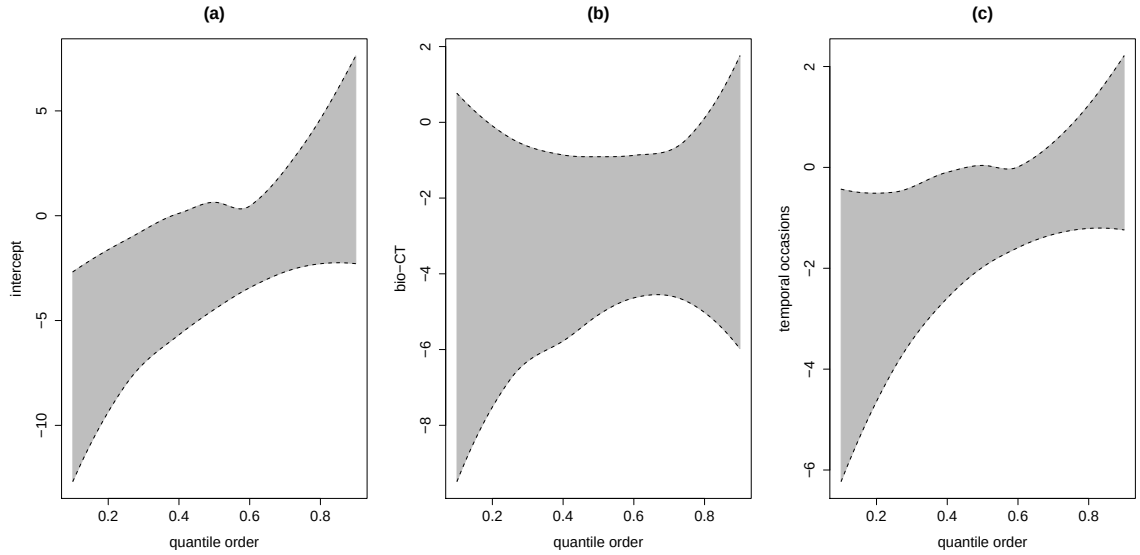

Figure 6: Parameter estimates and corresponding 95% confidence intervals estimates: (a) intercept, (b) BIO-CT treatment and (c) temporal occasion.

different procedures for data gathering. To standardize the procedures a common study protocol is usually adopted by participating centres. If the protocol has been properly planned and applied the centre heterogeneity is expected to be negligible. Hence, this result confirms the quality of the multi-centric design implemented for collecting the data used in this paper.

For this reason we have also fitted the two-levels MQRE (Tzavidis et al., 2015) and the LQMM (Geraci and Bottai, 2014). Table 6 reports the estimated parameter for the two-levels MQRE: the magnitude and the sign of the regression coefficients don't change respect the three-levels MQRE, as well as, the values of the estimated variance parameters for the patient and temporal occasion levels. Table 7 shows the estimates of the regression coefficients and variance component for the two-levels LQMM (random intercept). Note that the `lqmm` package does not report the standard errors of the variance components for the LQMM. It is evident that there are not differences in the sign of the regression coefficients respect the MQRE models. This means that the effect of the covariates do not change between the MQRE and LQMM models at different quantiles. The major difference is in the magnitude of the estimated variance parameter. The estimate of the patient effect level by LQMM shows higher values of variability respect that given by MQRE especially at the tail of the conditional distribution. The MQRE allows the estimation of a two-level and three-levels random intercepts model but it does not allow for

more complex correlation structures, including random coefficient models. In contrast, the `lqmm` in `R` allows for the specification of both random intercepts and random coefficients for the two-levels models. For this reason, we have used the LQMM with random intercepts specified at the level of the patient and random slopes (coefficients) specified for time. Random intercepts imply a uniform (exchangeable) correlation structure whereas random slopes allow the correlation structure to depend on time, which may offer a more realistic structure for repeated measures data. Table 8 reports the estimated parameter for this model. The estimates are comparable in terms of magnitude with those obtained by the other models. The estimated variance component for the time does not show high variability between temporal occasions. Considering that allowing for random slopes in quantile random effects models is complex and can potentially result in convergence problems when fitting the model and that quantile models with a random intercepts specification have a correlation structure that is simpler to estimate whilst allowing for modelling the entire conditional distribution of the outcome, the two and three-levels random intercepts MQRE and two-levels LQMM could be appropriate for this application.

6 Discussion and conclusion

The estimates reported in Table 5 and depicted in Figure 6 indicate that the bio-CT effect on PSDS score changed quite substantially at different quantiles. At the lower end of the distribution of PSDS the effect is strongly negative meaning that bio-CT reduces the PSDS score of patients compared to the standard treatment (CT). At the upper part of the distribution of PSDS, the effect of bio-CT declined becoming negligible and not statistically significant. This offers some evidence that patients can in fact benefit from the bio-CT (although this is not statistically significant). Hence, for those patients, the experiment regimen can be highly recommended, given the positive effect that this also has on the survival. In the case that one had considered only a conventional random effects linear model, this picture we get out of the quantile random effects model would have been completely lost. The treatment effect on the expected value of PSDS is negative (-3.29, sd=1.73) and not significant at the 5% level (p-value=0.0583).

Similarly we also found that the effect of time changed at different quantiles. The exposure to treatments reduced the PSDS score much more at the lower quantiles than at the highest ones, where this effect is not significant. This means that those patients who are in a better HRQOL state tolerated the treatment reasonably well whereas those who are in a poorer state suffered from the exposure to the therapy. In addition this information is completely lost, if the data was analyzed using a standard hierarchical linear random effects model. In this case one can only conclude that the HRQOL tends to decline on average as the study goes on with the slope being -1.38 (sd=0.60, p-value 0.0214).

As mentioned above the MQRE allows for the estimation of a two and three-levels random intercepts model but it does not allow for more complex correlation structures, including random coefficient models. To evaluate the stability of our results as compared

to other model allowing for more complicate correlation structures, we have also estimated a LQMM two-level random slope model via the `lqmm` package of R. Although the MQRE and LQMM results are not directly comparable as these models are targeting different location parameters, both models attempt to model location parameters that are associated with the same part of the conditional distribution of HRQOL scores. We found that the bio-CT coefficients obtained from `lqmm` have the same sign as the MQRE hence confirming our results.

In a multi-centre longitudinal trial, heterogeneity is often an issue and participating centres usually resort to a common study protocol to standardize the procedures and identifying eligible patients. Despite this, a large variability is often observed, hence, one of the goal of this paper was to evaluate if a centre effect impacts on the outcome distribution resorting to a MQRE three-level model.

Figure 7 shows the estimated intra-class correlation (ICC) of the MQRE at different quantiles. It is interesting to note that the ICC follows an inverted U curve. That is, both intra-centre and intra-patient correlation are higher in the middle of the outcome distribution where, proportionally, the within variability is smaller. Differences between centres and between patients, therefore, might play a less important role in explaining the total variability in below- and above-the-average quality of life of patients.

Via the MQRE analysis reported in the previous section we have demonstrated that centre heterogeneity is not an issue for the data at hand since it has been found negligible at all quantiles suggesting that the implemented protocol has succeeded in standardizing the data collection amongst centres. Nonetheless the methodology proposed in the present paper allows one to investigate centre heterogeneity in depth. Hence it worths to consider further this point.

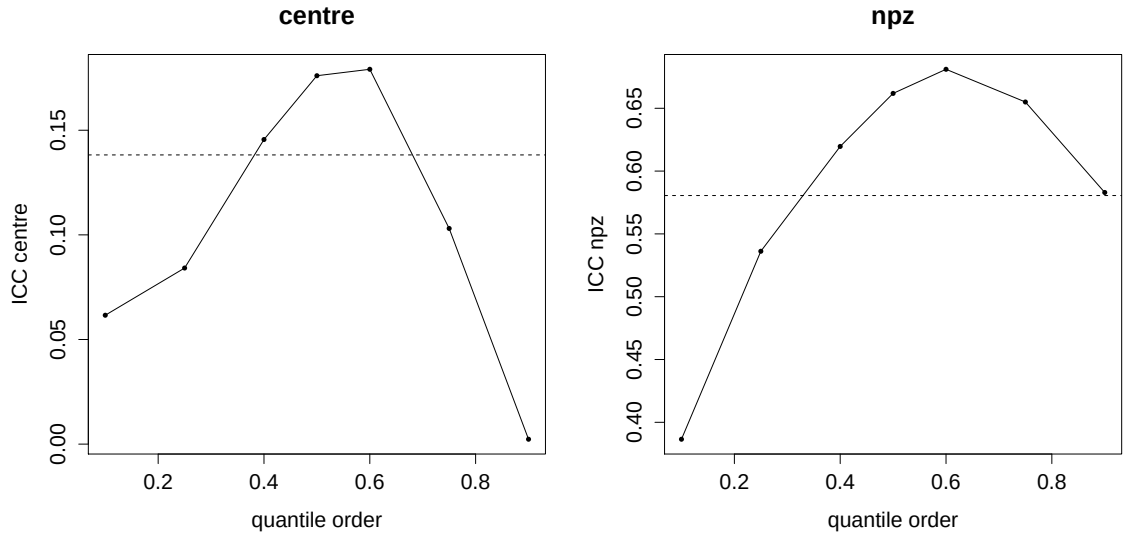

Figure 7: Estimated intra-class correlation (ICC) of the MQRE at quantiles.

Figure 8 depicts the ranks of the centre effect estimated at a low quantile order (0.10) plotted versus the ranks of the same effect estimated at a high quantile order (0.90). If the centre effect were the same at both quantiles this would imply that the points were

aligned on the first bisector which is also reported in the graph as a dashed line.

On the contrary from our analysis it turns out that some centres had a big effect (low rank) on the HRQOL at the higher quantile whereas they ranked very highly at the lower quantile. Motivations for this may be various. One may speculate that the recruitment of patients might have been different, at least to some extent, for different hospitals. Physicians, for instance, might have cared differently in administering questionnaires in some centres. One may argue that it is easier to obtain reliable information from patients in a reasonably good health, whereas this can be more problematic from those in poor health, particularly if and where appropriate staff, i.e. psychologists or ad hoc prepared hospital nurses, are not available as often occurs in smaller centres. This could result in a positive effect on HRQOL of some centres at the lower tail and may disappear at the higher quantiles where the performance of the hospitals can be sometimes reversed. This would also be consistent with the largest centre variability at the highest quantile modelled observed above.

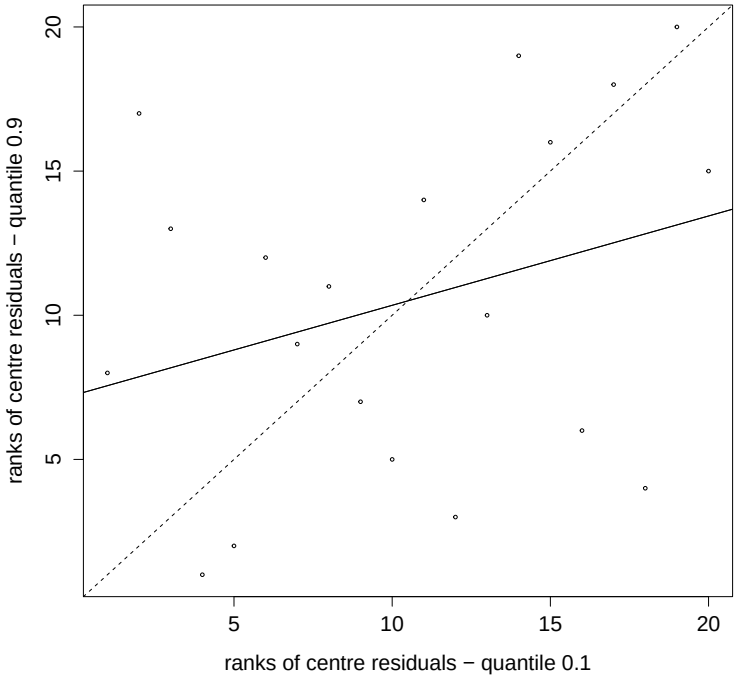

Figure 8: Plot of centre random effect ranks (MQRE for $q = 0.10, 0.90$).

This seems somehow confirmed by observing that the four centres which lay closer to the first bisector (dashed line) of Figure 8, i.e. those which preserved the ranks performing in a similar manner both at the lower and at the upper quantile, are in fact leading centres of the trial considered in this paper, all of them having a reasonably large number of enrolled patients (at least 10). Looking at the estimated residuals of centres of the M-quantile regression of order 0.9, one may also notice that all of the centres which enrolled at least 10 patients (the larger ones) show estimates below the average of the estimated residuals. This proportion drops to 15.8% for centres with less than 10 patients. On the contrary, while exactly a half of the bigger centres are above the average of esti-

mated residuals when the M-quantile regression of order 0.1 is considered, the proportion of estimated residuals above the average drops to 36.8% (less than what happened for quantile 0.9) for smaller centres. This consideration seems to be consistent with a sort of enrollment bias that may have occurred in some of the participating centres.

This also suggests that a plot like the one reported in Figure 8 can also be usefully employed in ad interim analysis to point out potential anomalous behaviours or situations which may deserve some more in-depth ad hoc investigation. Finally, box-centile plots of the predicted random effects are shown in Figure 9.

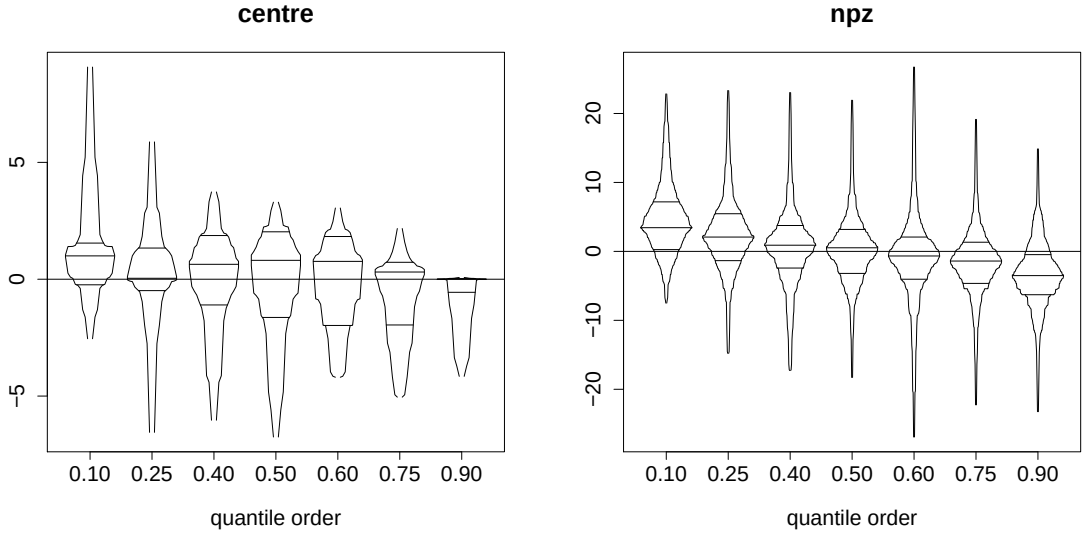

Figure 9: Box-centile plots of predicted centre and npz-specific random effects.

We can observe that the drop out in this study is not negligible as in many other longitudinal studies on HRQOL (see amongst others a recent work of le Cessie et al., 2009) particularly in chronic diseases where the survivorship is extremely short. In our dataset only a minor part withdrew from the study for reasons other than the trial design. In fact nearly all of them dropped out of the study due to disease progression hence the treatment was interrupted and their quality of life measurements were no longer collected. A usual way to compensate a disproportioned drop-out is via sampling weights as far as the drop out mechanism can be considered at random (Little and Rubin, 2002). Lipsitz et al. (1997) and Yi and He (2009) investigated the use of weighting within a generalized equation estimation framework in quantile and median regression respectively, an approach akin to what Robins et al. (1995) proposed for estimating mean regression. A part from these two remarkable pieces of work, no other attempts have been made to adjust for non response in fix effect quantile models. No papers at all have dealt with this issue in fix effect M-quantile modelling. Embedding weights in usual mixed model is a difficult task firstly addressed by Pfeffermann et al. (1998). Estimated weights may heavily affect the inference procedures both for the estimation of the model and for hypotheses testing. Weighting-adjusted inferential procedures in random effect M-quantile modelling have not been addressed so far to the best of our knowledge. Investigating drop out adjustment in fix as well as random effect M-quantile models is a challenging issue which was, however,

far beyond the aim of the present paper and it is a matter for future research.

Appendix A

The steps of the estimation algorithm as it follows:

1 Start by assuming that $(\sigma^2_{u_q}, \sigma^2_{\gamma_q}, \sigma^2_{\epsilon_q})$ are known.

2 Given these variance parameters, form the covariance matrix $\mathbf{V}_q$, and estimate $\boldsymbol{\beta}_q$ by solving the iterative equation

$$\boldsymbol{\beta}_q^{t+1} = \boldsymbol{\beta}_q^t + \{\mathbf{X}^T\mathbf{U}_q^{-1/2}\mathbf{H}_q(\boldsymbol{\beta}_q^t)\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{X}\}^{-1}\mathbf{X}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\},$$

where $\mathbf{H}_q(\boldsymbol{\beta}_q^t)$ is a diagonal matrix with its jth diagonal element $H_{ijkq} = \psi'_q(r_{ijkq}) = (\partial/\partial r_{ijkq})\psi_q(r_{ijkq})$.

3 Use the estimates of $\boldsymbol{\beta}_q$ to obtain estimates of the variance parameters. The estimates of the variance parameters are obtained by using fixed-point iterative method. This requires us to change the estimating equations (5) as:

$$\begin{aligned}
\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\left(\mathbf{D}\mathbf{D}^T \;\; \mathbf{Z}\mathbf{Z}^T \;\; \mathbf{I}_N\right)\begin{pmatrix}\sigma^2_{u_q}\\ \sigma^2_{\gamma_q}\\ \sigma^2_{\epsilon_q}\end{pmatrix}\right] &= 0\\
\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\left(\mathbf{D}\mathbf{D}^T \;\; \mathbf{Z}\mathbf{Z}^T \;\; \mathbf{I}_N\right)\begin{pmatrix}\sigma^2_{u_q}\\ \sigma^2_{\gamma_q}\\ \sigma^2_{\epsilon_q}\end{pmatrix}\right] &= 0\\
\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\} - K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\left(\mathbf{D}\mathbf{D}^T \;\; \mathbf{Z}\mathbf{Z}^T \;\; \mathbf{I}_N\right)\begin{pmatrix}\sigma^2_{u_q}\\ \sigma^2_{\gamma_q}\\ \sigma^2_{\epsilon_q}\end{pmatrix}\right] &= 0,
\end{aligned}$$

by replacing $\mathbf{V}_q$ by $\mathbf{V}_q = \sigma^2_{u_q}\mathbf{D}\mathbf{D}^T + \sigma^2_{\gamma_q}\mathbf{Z}\mathbf{Z}^T + \sigma^2_{\epsilon_q}\mathbf{I}_N$ and using $\mathbf{V}_q^{-1}\mathbf{V}_q = \mathbf{I}_N$. The fixed point algorithm of the estimating equations for the tth iteration can be presented as follows:

$$\begin{pmatrix}\sigma^{2(t+1)}_{u_q}\\ \sigma^{2(t+1)}_{\gamma_q}\\ \sigma^{2(t+1)}_{\epsilon_q}\end{pmatrix} = \left[\mathbf{A}\begin{pmatrix}\sigma^{2(t)}_{u_q}\\ \sigma^{2(t)}_{\gamma_q}\\ \sigma^{2(t)}_{\epsilon_q}\end{pmatrix}\right]^{-1} a\begin{pmatrix}\sigma^{2(t)}_{u_q}\\ \sigma^{2(t)}_{\gamma_q}\\ \sigma^{2(t)}_{\epsilon_q}\end{pmatrix}, \qquad \text{(A-1)}$$

where

$$\mathbf{A}\begin{pmatrix}\sigma^2_{u_q}\\ \sigma^2_{\gamma_q}\\ \sigma^2_{\epsilon_q}\end{pmatrix} = \begin{pmatrix} K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{I}_N\right]\\ K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{I}_n\right]\\ K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\right] & K_{2q}tr\left[\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\mathbf{I}_N\right]\end{pmatrix}$$

and

$$a\begin{pmatrix}\sigma^2_{u_q}\\ \sigma^2_{\gamma_q}\\ \sigma^2_{\epsilon_q}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{D}\mathbf{D}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\}\\ \frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{Z}\mathbf{Z}^T\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\}\\ \frac{1}{2}\psi_q\{\mathbf{r}_q\}^T\mathbf{U}_q^{1/2}\mathbf{V}_q^{-1}\mathbf{I}_N\mathbf{V}_q^{-1}\mathbf{U}_q^{1/2}\psi_q\{\mathbf{r}_q\}\end{pmatrix}.$$

Iterative equation (A-1) is more stable than the Newton-Raphson method and it typically converges in 10 to 15 steps. Like any other iterative algorithm, the fixed-point algorithm requires initial values for the parameters. As a result, suggesting a well-defined starting value for the variance parameters could facilitate the procedure.

4 Iterate steps 2 and 3 until convergence.

5 At convergence, estimates of the random effects at qth quantile fit are obtained by solving the following estimating equations with respect to $\mathbf{u}_q$ and $\boldsymbol{\gamma}_q$

$$\mathbf{D}^T\boldsymbol{\Sigma}_{\epsilon_q}^{-1/2}\psi_q\{\boldsymbol{\Sigma}_{\epsilon_q}^{-1/2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}_q-\mathbf{D}\mathbf{u}_q-\mathbf{Z}\boldsymbol{\gamma}_q)\}-\boldsymbol{\Sigma}_{u_q}^{-1/2}\psi_q\{\boldsymbol{\Sigma}_{u_q}^{-1/2}\mathbf{u}_q\}=\mathbf{0}. \quad \text{(A-2)}$$

$$\mathbf{Z}^T\boldsymbol{\Sigma}_{\epsilon_q}^{-1/2}\psi_q\{\boldsymbol{\Sigma}_{\epsilon_q}^{-1/2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}_q-\mathbf{D}\mathbf{u}_q-\mathbf{Z}\boldsymbol{\gamma}_q)\}-\boldsymbol{\Sigma}_{\gamma_q}^{-1/2}\psi_q\{\boldsymbol{\Sigma}_{\gamma_q}^{-1/2}\boldsymbol{\gamma}_q\}=\mathbf{0}. \quad \text{(A-3)}$$

As can be seen from the steps of the estimation algorithm, estimates of the random effects are obtained at convergence, i.e. we start by first estimating the fixed effects and the variance parameters and then given robust estimates of the fixed effects and of the variance parameters we estimate the random effects. The reason for this, as also pointed out by Sinha and Rao (2009), is that estimates of the variance parameters that depend on the estimated random effects are not statistically efficient.

References

Anderson, T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. *Annals of Statistics 1 (1)*, 135–141.

Austin, P. C., J. V. Tu, P. A. Daly, and D. A. Alter (2005). The use of quantile regression in health care research: a case study examining gender differences in the timeliness of thrombolytic therapy. *Statistics in Medicine 24 (5)*, 791–816.

Bottai, M., B. Cai, and R. E. McKeown (2010). Logistic quantile regression for bounded outcomes. *Statistics in Medicine 29 (2)*, 309–317.

Breckling, J. and R. Chambers (1988). M-quantiles. *Biometrika 75 (4)*, 761–771.

Broccoli, S., G. Cavrini, and M. Zoli (2005). A quantile regression approach to the analysis of the quality of life determinants in the elderly. *Statistica 65 (4)*, 419–436.

Carey, V. J., F. H. Yong, L. M. Frenkel, and R. M. McKinney (2004). Growth velocity assessment in paediatric aids: smoothing, penalized quantile regression and the definition of growth failure. *Statistics in Medicine 23 (3)*, 509–526.

Chiarion-Sileni, V., P. D. Bianco, G. D. Salvo, G. Re, A. Romanini, R. Labianca, R. Nortilli, E. Corgna, M. Palma, G. Presti, and R. Ridolfi (2003). Quality of life evaluation in a randomised trial of chemotherapy versus bio-chemotherapy in advanced melanoma patients. *European Journal of Cancer 39 (11)*, 1577–1585.

Fairclough, D. (2002). *Design and Analysis of Quality of Life Studies in Clinical Trials*. Chapman & Hall/CRC Interdisciplinary Statistics. Taylor & Francis.

Fellner, W. H. (1986). Robust estimation of variance components. *Technometrics 28 (1)*, 51–60.

Geraci, M. and M. Bottai (2007). Quantile regression for longitudinal data using the asymmetric laplace distribution. *Biostatistics 8 (1)*, 140–154.

Geraci, M. and M. Bottai (2014). Linear quantile mixed models. *Statistics and Computing 24 (3)*, 461–479.

Goldstein, H. (2003). *Multilevel Statistical Models*. New York: Wiley.

Guthrie, L., E. Oken, J. Sterne, M. Gillman, R. Patel, K. Vilchuck, N. Bogdanovich, M. Kramer, and R. Martin (2012). Ongoing monitoring of data clustering in multicenter studies. *BMC Medical Research Methodology 12*.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association 72 (358)*, 320–338.

Huber, P. (1981). *Robust Statistics*. New York: Wiley.

Huggins, R. M. (1993). A robust approach to the analysis of repeated measures. *Biometrics 49 (3)*, 715–720.

Huggins, R. M. and D. Z. Loesch (1998). On the analysis of mixed longitudinal growth data. *Biometrics 54 (2)*, 583–595.

Jones, M. C. (1994). Expectiles and m-quantiles are quantiles. *Statistics and Probability Letters 20*, 149–153.

Koenker, R. (2005). *Quantile Regression*. New York: Cambridge University Press.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

le Cessie, S., E. G. E. de Vries, C. Buijs, and W. J. Post (2009). Analyzing longitudinal data with patients in different disease states during follow-up and death as final state. *Statistics in Medicine 28 (30)*, 3829–3843.

Li, Y., B. I. Graubard, and E. L. Korn (2010). Application of nonparametric quantile regression to body mass index percentile curves from survey data. *Statistics in Medicine 29 (5)*, 558–572.

Lipsitz, S. R., G. M. Fitzmaurice, G. Molenberghs, and L. P. Zhao (1997). Quantile regression methods for longitudinal data with drop-outs: Application to cd4 cell counts of patients infected with the human immunodeficiency virus. *Journal of the Royal Statistical Society: Series C 46 (4)*, 463–476.

Little, R. J. A. and D. B. Rubin (2002). *Statistical Analysis with Missing Data*. New York, NY, USA: John Wiley & Sons, Inc.

Localio, A. R., J. A. Berlin, T. R. Ten Have, and S. E. Kimmel (2001). Adjustments for center in multicenter studies: An overview. *Annals of Internal Medicine 135 (2)*, 112–123.

Newey, W. K. and J. L. Powell (1987). Asymmetric least squares estimation and testing. *Econometrica 55 (4)*, 819–847.

Nicholson, W., R. Setse, F. Hill-Briggs, L. Cooper, D. Strobino, and N. Powe (2006). Depressive symptoms and health-related quality of life in early pregnancy. *Obstetrics Gynecology 107 (4)*, 798–806.

Pfeffermann, D., C. J. Skinner, D. J. Holmes, H. Goldstein, and J. Rasbash (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society: Series B 60 (1)*, 23–40.

Pourhoseingholi, M. A., A. Safaee, B. Moghimi-Dehkordi, B. Zeighami, S. Faghihzadeh, H. R. Tabatabaee, and A. Pourhoseingholi (2008). Quality of life in breast cancer patients - a quantile regression analysis. *Asian Pacific Journal of Cancer Prevention 9 (3)*, 487–490.

Richardson, A. M. (1997). Bounded influence estimation in the mixed linear model. *Journal of the American Statistical Association 92 (437)*, 154–161.

Richardson, A. M. and A. H. Welsh (1995). Robust restricted maximum likelihood in mixed linear models. *Biometrics 51 (4)*, 1429–1439.

Ridolfi, R., V. Chiarion-Sileni, M. Guida, A. Romanini, R. Labianca, A. Freschi, G. L. Re, R. Nortilli, S. Brugnara, P. Vitali, and O. Nanni (2002). Cisplatin, dacarbazine with or without subcutaneous interleukin-2, and interferon alpha-2b in advanced melanoma outpatients: results from an italian multicenter phase iii randomized clinical trial. *Journal of Clinical Oncology*, 1600–1607.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association 90 (429)*, 106–121.

Sakamoto, J., Y. Matsuyama, and Y. Ohashi (1999). An analysis of institutional effects in a multicenter cancer clinical trial: is it also plausible from the clinicians' point of view? *Japanese Journal of Clinical Oncology 29 (8)*, 403–405.

Sinha, S. K. and J. N. K. Rao (2009). Robust small area estimation. *The Canadian Journal of Statistics 37 (3)*, 381–399.

Tzavidis, N., N. Salvati, T. Schmid, E. Flouri, and E. Midouhas (2015). Longitudinal analysis of the strengths and difficulties questionnaire scores of the millennium cohort study children in england using m-quantile random effects regression. *Journal of the Royal Statistical Society: Series A*, forthcoming.

Wei, Y., A. Pere, R. Koenker, and X. He (2006). Quantile regression methods for reference growth charts. *Statistics in Medicine 25 (8)*, 1369–1382.

Yamaguchi, T., Y. Ohashi, and Y. Matsuyama (2002). Proportional hazards models with random effects to examine centre effects in multicentre cancer clinical trials. *Statistical Methods in Medical Research 11 (3)*, 221–236.

Yi, G. Y. and W. He (2009). Median regression models for longitudinal data with dropouts. *Biometrics 65 (2)*, 618–625.

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business and Economics - Freie Universität Berlin

2015 erschienen:

2015/1 GÖRLITZ, Katja und Christina GRAVERT
The effects of increasing the standards of the high school curriculum on school dropout
Economics

2015/2 BÖNKE, Timm und Clive WERDT
Charitable giving and its persistent and transitory reactions to changes in tax incentives: evidence from the German Taxpayer Panel
Economics

2015/3 WERDT, Clive
What drives tax refund maximization from inter-temporal loss usage? Evidence from the German Taxpayer Panel
Economics

2015/4 FOSSEN, Frank M. und Johannes KÖNIG
Public health insurance and entry into self-employment
Economics

2015/5 WERDT, Clive
The elasticity of taxable income for Germany and its sensitivity to the appropriate model
Economics

2015/6 NIKODINOSKA, Dragana und Carsten SCHRÖDER
On the Emissions-Inequality Trade-off in Energy Taxation: Evidence on the German Car Fuel Tax
Economics

2015/7 GROß, Marcus; Ulrich RENDTEL; Timo SCHMID; Sebastian SCHMON und Nikos TZAVIDIS
Estimating the density of ethnic minorities and aged people in Berlin: Multivariate kernel density estimation applied to sensitive geo-referenced administrative data protected via measurement error
Economics

2015/8 SCHMID, Timo; Nikos TZAVIDIS; Ralf MÜNNICH und Ray CHAMBERS
Outlier robust small area estimation under spatial correlation
Economics

2015/9 GÖRLITZ, Katja und Marcus TAMM
Parenthood and risk preferences
Economics

2015/10 BÖNKE, Timm; Giacomo CORNEO und Christian WESTERMEIER
Erbschaft und Eigenleistung im Vermögen der Deutschen: eine Verteilungsanalyse
Economics

2015/11 GÖRLITZ, Katja und Marcus TAMM
The pecuniary and non-pecuinary returns to voucher-financed training
Economics

2015/12 CORNEO, Giacomo
Volkswirtschaftliche Bewertung öffentlicher Investitionen
Economics

2015/13 GÖRLITZ, Katja und Christina Gravert
The effects of a high school curriculum reform on university enrollment and the choice of college major
Economics

2015/14 BÖNKE, Timm und Carsten SCHRÖDER
European-wide inequality in times of the financial crisis
Economics

2015/15 BÖNKE, Timm; Beate JOACHIMSEN und Carsten SCHRÖDER
Fiscal federalism and tax enforcement
Economics

2015/16 DEMMER, Matthias
Improving Profitability Forecasts with Information on Earnings Quality
FACTS

2015/17 HAAN, Peter und Victoria PROWSE
Optimal Social Assistance and Unemployment Insurance in a Life-cycle Model of Family Labor Supply and Savings
Economics

2015/18 CORNEO, Giacomo, Carsten SCHRÖDER und Johannes KÖNIG
Distributional Effects of Subsidizing Retirement Savings Accounts: Evidence from Germany
Economics