

# 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,<sup>1</sup> Paola Del Bianco,<sup>2</sup> Nicola Salvati,<sup>3</sup> Timo Schmid<sup>4</sup> and Nikos Tzavidis<sup>5</sup>

Statistical Methods in Medical Research  
2018, Vol. 27(2) 549–563

© The Author(s) 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0962280216636651

journals.sagepub.com/home/smm



## 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 quantile and M-quantile regression to achieve this goal. We applied the methodologies 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, quantile regression, robust estimation, Rotterdam Symptom Checklist, Phase III study

## 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, 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 are 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 enroll 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

<sup>1</sup>Dipartimento di Economia, Metodi Quantitativi e Strategie d'Impresa, Università di Milano – Bicocca, Milan, Italy

<sup>2</sup>Clinical Trials and Biostatistics Unit, Istituto Oncologico Veneto IOV-IRCCS, Padova, Italy

<sup>3</sup>Dipartimento di Economia e Management, Università di Pisa, Pisa, Italy

<sup>4</sup>Institute of Statistics and Econometrics, Freie Universität Berlin, Berlin, Germany

<sup>5</sup>Department of Social Statistics and Demography, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, UK

## Corresponding author:

Nicola Salvati, Dipartimento di Economia e Management, Università di Pisa, Via Ridolfi, 10-I56124 Pisa, Italy.

Email: nicola.salvati@unipi.it

inter-centre variability as a result of differences in applying the study protocol procedures.<sup>1,2</sup> 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.<sup>3,4</sup>

Centre effects can therefore be more conveniently specified by using a multilevel (random effects) modelling approach<sup>5</sup> 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.<sup>6</sup>

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 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<sup>7</sup> and thoroughly described by Koenker.<sup>8</sup> Since then, quantile regression has been applied in many and different areas of research. Recently, it has been successfully employed in medical applications.<sup>9-14</sup> Only few applications of quantile regression to modelling HRQOL currently exist in literature,<sup>15-17</sup> but the studies do not take into account the hierarchical structure of the data. Recently, Geraci and Bottai<sup>12,18</sup> propose a conditional two-level 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,<sup>19</sup> which integrates the concepts of quantile regression and expectile regression<sup>20</sup> 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  $MQ_{y_i}(q|\mathbf{X}; \psi)$  of the estimating equation  $\int \psi_q\{y - MQ_{y_i}(q|\mathbf{X}; \psi)\}f(y|\mathbf{X})dy = 0$ , where  $\psi_q$  denotes an asymmetric influence function, which is the derivative of an asymmetric loss function  $\rho_q$ . 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

$$\sum_{i=1}^n \rho_q\{r_{iq}\} \quad (1)$$

where  $r_{iq} = \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}_q}{\sigma}$ ,  $\sigma$  is a scale parameter,  $\rho_q\{r_{iq}\} = 2\rho\{r_{iq}\}[qI(r_{iq} > 0) + (1 - q)I(r_{iq} \leq 0)]$  and  $I(\cdot)$  is the indicator function. Different set of regression parameters can be defined for each value of  $q$ . In particular, by varying the specifications of the asymmetric loss function  $\rho$  we obtain the expectile, M-quantile and quantile regression models as special cases. When  $\rho$  is the square loss function,  $\rho_q\{r_{iq}\} = 2r_{iq}^2[qI(r_{iq} > 0) + (1 - q)I(r_{iq} \leq 0)]$ , we obtain the linear expectile regression model if  $q \neq 0.5$ <sup>20</sup> and the standard linear regression model if  $q = 0.5$  ( $\rho_{0.5}\{r_{i0.5}\} = r_{i0.5}^2$ ). When  $\rho$  is the loss function described by Koenker and Bassett,<sup>7</sup>  $\rho_q\{r_{iq}\} = |r_{iq}|[qI(r_{iq} > 0) + (1 - q)I(r_{iq} \leq 0)]$ , we

obtain the linear quantile regression. Throughout this paper we take the linear M-quantile regression model to be defined by using as  $\rho$  the Huber loss function.<sup>21</sup> In particular, the tilted version of the popular Huber loss function can be written as

$$\rho_q\{r_{iq}\} = 2 \begin{cases} (c|r_{iq}| - c^2/2)|q - I(r_{iq} \leq 0)| & |r_{iq}| > c \\ r_{iq}^2/2|q - I(r_{iq} \leq 0)| & |r_{iq}| \leq c \end{cases} \quad (2)$$

where  $c$  is a cutoff constant. M-quantile regression models allow us to trade robustness for efficiency by properly tuning the constant  $c$ : robustness is increased as  $c$  decreases, while efficiency is increased as  $c$  increases.

Tzavidis et al.<sup>22</sup> extended M-quantile regression to include random effects (MQRE-2L) to account for a two-level hierarchical structure in the data and used maximum likelihood (ML) to estimate the parameters of the model.

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-level models). In order to account for the hierarchical nature of the data we include two random effects to capture the between centre variability and to account simultaneously for unobserved individual heterogeneity, i.e. the variability due to repeated measurements on the same patient. Therefore, we extend the MQRE-2L to a three-level random effects model (MQRE-3L). From an applied point of view, the paper presents a methodology (MQRE-3L) to investigate potential centre heterogeneity caused as a result of differences in applying the study protocol procedures. The conditional distribution of HRQOL in advanced melanoma patients is modelled by LQMM, MQRE-2L and MQRE-3L using the data of the randomized multi-centre clinical trial.

Quantiles have a more intuitive interpretation than M-quantiles even if they target essentially the same part of the distribution of interest.<sup>23</sup> 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.

The paper is structured as follows. The data are introduced in Section 2. In Section 3, we present the results from the application of two and three-level M-quantile random effects regression model and quantile random effects regression to the HRQOL data. Therefore, we introduce in Section 3.2 the extension to three-level M-quantile random effects regression. The results of the application are discussed and concluding remarks are presented in Section 4.

## 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<sup>2</sup> and DTIC 800 mg/m<sup>2</sup> (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 HRQOL analysis are reported elsewhere.<sup>24,25</sup>

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.

### 3 M-quantile modelling of HRQOL

In this section, the conditional distribution of HRQOL in advanced melanoma patients is modelled 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 modelled 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.

As far as the random component at the patient level is concerned it can be noted that modeling the change from the baseline adjusts for the within patients heterogeneity due to baseline differences. However, it is expected that other sources of variability can act at the individual level such as the tolerance of patients to the drug, the emotional support they receive and the health status related to other aspects than those under study. All these sources of variability are clearly not adjusted by modeling the change from the baseline. These factors can strongly impact on the perceived quality of life inducing correlation between the individual HRQOL measurements and motivate the random component at the patient level in the first place.

There is also a substantive and policy-related reason as to why one may wish to test for a centre effect. The study related to the implementation of a clinical trial protocol and it is important to test that the protocol has been implemented consistently across different centres in the trial. Lack of a harmonized implementation of such a protocol will raise serious ethical issues. Hence, a three-level model is needed for testing for such an effect.

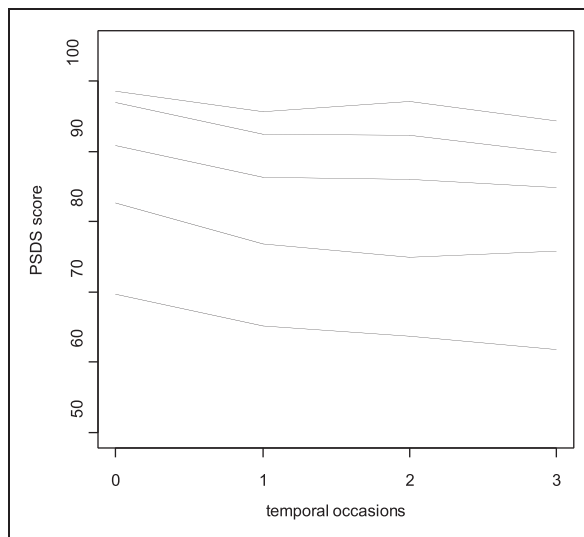
The LQMM and the MQRE-2L do not allow to capture simultaneously the between centre variability and the individual heterogeneity due to repeated measurements by random effects. For this reason we extend in Section 3.2

**Table 1.** Summary statistics of PSDS stratified by treatment.

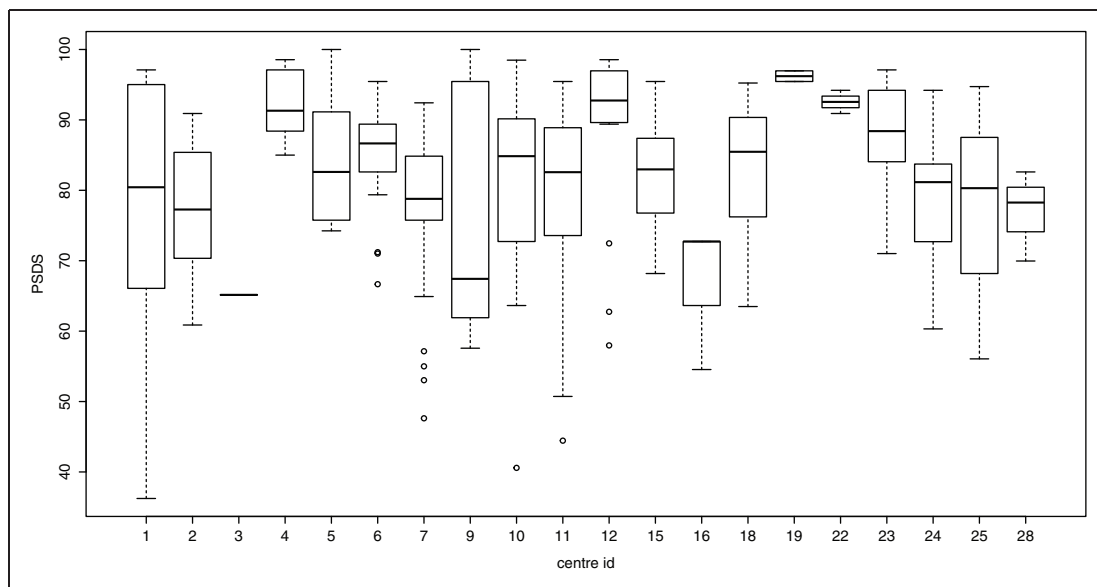
Order of sample quantiles								
Treatment	N. Obs	0.10	0.25	0.50	0.75	0.90	Mean	SD
CT	264	68.70	78.26	87.88	93.94	97.10	84.73	11.82
Bio-CT	244	64.26	75.76	86.66	92.75	97.10	83.17	12.72

**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	SD
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



**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.

the MQRE-2L to a three-level MQRE. Although the extension of the two-level model given in Tzavidis et al.<sup>22</sup> to three-level MQRE is conceptually straightforward, the description of the MQRE-3L requires the introduction of some technical details.

### 3.1 Preliminary data analyses

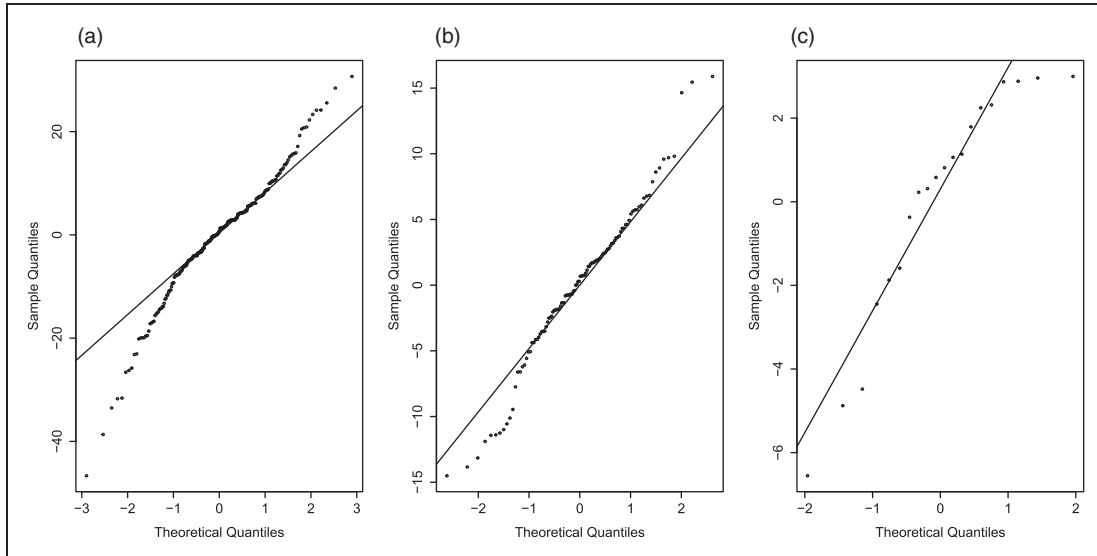
Before showing the results from the study we present some preliminary data analysis. We applied a model with the following three-level hierarchical structure because in the multi-centre clinical trial data set, occasions (repeated measures denoted by  $k$  – level 1) are nested within patients ( $npz$ , denoted by  $j$  – level 2), and patients are nested within *centres* (denoted by  $i$  – level 3):

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

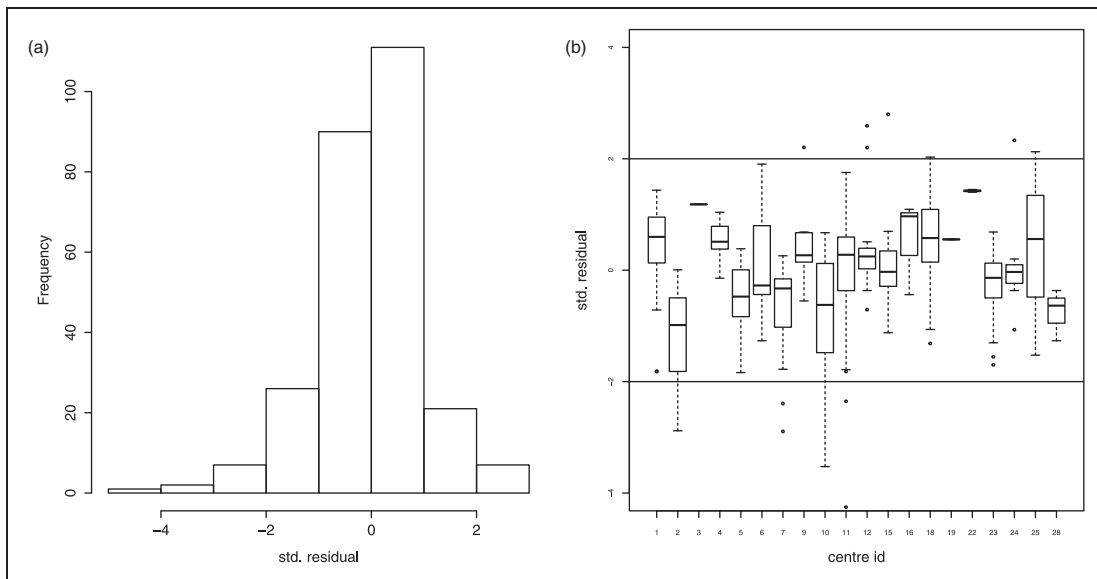
$$i = 1, \dots, m = 23, j = 1, \dots, n_i, k = 1, \dots, t_{ij}$$

where  $y_{ijk}$  is the change of the HRQOL score at each time point from the baseline,  $\sum_{i=1}^m n_i = 137$  and  $\sum_{i=1}^m \sum_{j=1}^{n_i} t_{ij} = 508$ . 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. 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).

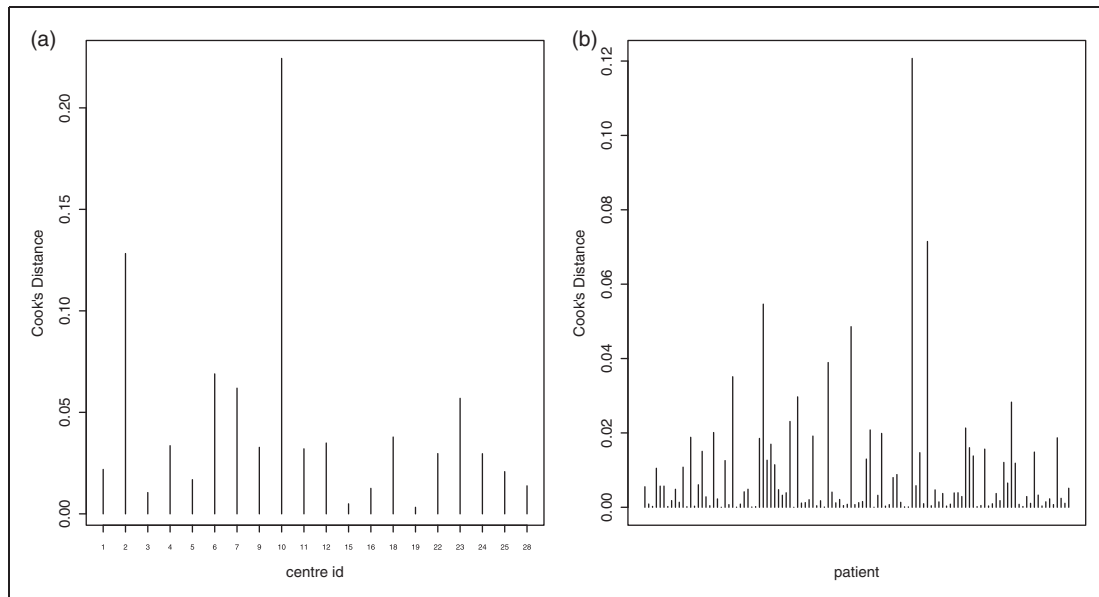
Figure 4 shows two plots of standardized mixed model residuals. The histogram, plot (a), depicts a skewed distribution of the residuals. This is confirmed by plot (b) which shows the distribution of standardized mixed model residuals by centre: some centres have many positive residuals, whereas others have many negative



**Figure 3.** Normal probability plots of level 1 (a) level 2 (b) and level 3 residuals (c) derived by fitting a three-level 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).



**Figure 5.** Cook's distance by centre (a) and by patient (b).

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.

### 3.2 Three-level M-quantile random effects regression

In what follows, we present a three-level M-quantile random effects regression. A conventional three-level random effects model in matrix form is described by

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

where  $\mathbf{y}$  denotes the values of the variable of interest  $y$ ,  $\mathbf{X}$  is a  $N \times p$  matrix of auxiliary variables,  $\mathbf{Z}$  and  $\mathbf{D}$  are auxiliary contextual matrices. We assume that  $\mathbf{X}$  contains 1 as its first component and  $\mathbf{Z}$  and  $\mathbf{D}$  are incidence matrices which specify the random intercepts specification of model. Here,  $\boldsymbol{\beta}$  is a  $p$  vector of fixed effect parameters and  $\mathbf{u}$ ,  $\boldsymbol{\gamma}$ ,  $\boldsymbol{\varepsilon}$  represents the random effects (centre specific, patient and the occasion random effects). Model (3) conventionally assumes that  $\mathbf{u} \sim N(\mathbf{0}, \sigma_u^2 \mathbf{I}_m)$ ,  $\boldsymbol{\gamma} \sim N(\mathbf{0}, \sigma_\gamma^2 \mathbf{I}_n)$ ,  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_N)$  and mutually independent.  $\mathbf{I}_g$  is an identity matrix of size  $g$ .

ML estimation based on the marginal distribution of  $y$  is widely used for estimating the unknown parameters of model (3).<sup>26</sup> In particular, estimates of  $\boldsymbol{\beta}$ ,  $\sigma_u^2$ ,  $\sigma_\gamma^2$  and  $\sigma_\varepsilon^2$  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.<sup>5</sup> Estimates of the random effects are then obtained by using the ML estimates of the fixed effects and the variance components.

Data may contain outliers that invalidate the Gaussian assumptions as it occurs in the case study analyzed in this paper. In such a case, the estimated model parameters under equation (3) will be biased and inefficient.<sup>27</sup> A number of papers<sup>27–30</sup> 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.

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<sup>22</sup> 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. The main differences with respect to the MQRE-2L in Tzavidis et al.<sup>22</sup> are: (1) an

additional equation for estimating the variance component due to the presence of centre effect and (2) a more complex variance covariance-matrix of  $y$ . According to Sinha and Rao<sup>31</sup> and Tzavidis et al.,<sup>22</sup> we can extend the idea of asymmetric weighting of residuals to hierarchical data by defining modified estimating equations obtained by changing the estimating equations of the robust ML proposal II by Richardson and Welsh<sup>27</sup> as follows

$$\begin{aligned} \mathbf{X}^T \mathbf{V}_q^{-1} \mathbf{U}_q^{1/2} \psi_q\{\mathbf{r}_q\} &= \mathbf{0} \\ \frac{1}{2} \psi_q\{\mathbf{r}_q\}^T \mathbf{U}_q^{1/2} \mathbf{V}_q^{-1} \mathbf{D} \mathbf{D}^T \mathbf{V}_q^{-1} \mathbf{U}_q^{1/2} \psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2} \text{tr}[\mathbf{V}_q^{-1} \mathbf{D} \mathbf{D}^T] &= 0 \\ \frac{1}{2} \psi_q\{\mathbf{r}_q\}^T \mathbf{U}_q^{1/2} \mathbf{V}_q^{-1} \mathbf{Z} \mathbf{Z}^T \mathbf{V}_q^{-1} \mathbf{U}_q^{1/2} \psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2} \text{tr}[\mathbf{V}_q^{-1} \mathbf{Z} \mathbf{Z}^T] &= 0 \\ \frac{1}{2} \psi_q\{\mathbf{r}_q\}^T \mathbf{U}_q^{1/2} \mathbf{V}_q^{-1} \mathbf{V}_q^{-1} \mathbf{U}_q^{1/2} \psi_q\{\mathbf{r}_q\} - \frac{K_{2q}}{2} \text{tr}[\mathbf{V}_q^{-1}] &= 0 \end{aligned} \quad (4)$$

$$(5)$$

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  $\rho_q$ . Here,  $\mathbf{V}_q = \mathbf{D}\boldsymbol{\Sigma}_{u_q}\mathbf{D}^T + \mathbf{Z}\boldsymbol{\Sigma}_{\gamma_q}\mathbf{Z}^T + \boldsymbol{\Sigma}_{\epsilon_q}$ ,  $\boldsymbol{\Sigma}_{u_q} = \sigma_{u_q}^2 \mathbf{I}_m$ ,  $\boldsymbol{\Sigma}_{\gamma_q} = \sigma_{\gamma_q}^2 \mathbf{I}_n$ ,  $\boldsymbol{\Sigma}_{\epsilon_q} = \sigma_{\epsilon_q}^2 \mathbf{I}_N$ ,  $\sigma_{u_q}$ ,  $\sigma_{\gamma_q}$  and  $\sigma_{\epsilon_q}$  are the quantile-specific variance parameters, and  $\boldsymbol{\beta}_q$  is the  $p \times 1$  vector of M-quantile regression coefficients. Finally,  $K_{2q}$  is  $E[\psi_q(\varepsilon)\psi_q(\varepsilon)^T]$  with  $\varepsilon \sim N(\mathbf{0}, \mathbf{I}_N)$ .

With equation (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,<sup>21</sup> to M-quantile regression for hierarchically structured data. ML equations of model (3) and their robust version are a special case of the estimating equations in equation (5) for a specific choice of  $\rho_q$  and  $q$ . See Tzavidis et al.<sup>22</sup> for a detailed discussion.

As discussed in Jones,<sup>23</sup> 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 (4) and (5) can be solved iteratively to obtain estimators of  $\boldsymbol{\beta}_q$ ,  $\sigma_{u_q}$ ,  $\sigma_{\gamma_q}$  and  $\sigma_{\epsilon_q}$ . Here, we adopt Newton-Raphson algorithm for solving equation (4) and the fixed-point iterative method<sup>32</sup> for solving estimating equations (5). The steps of the estimation algorithm are outlined in the supplementary material. We have also worked on studying the algorithmic stability that is, ensuring that the convergence of the algorithm does not depend on the starting values. For the data we used in this paper, we studied the convergence and we can confirm that, when initialising the algorithm from different starting values, it always converged to the same point also for the extreme quantiles. However, there may be cases in real data applications where due to the sparsity of the data convergence is not stable at extreme quantile. Hence, users of the method should always test the convergence of the algorithm with their datasets to ensure that there are no convergence problems. The estimation of extreme quantiles is sometimes challenging when there is a small number of groups or when groups contain a small number of observations. However, in this case, an alternative estimation approach that can potentially lead to more efficient estimates of the variance components is the restricted ML approach.<sup>27,33</sup>

A function that fits the three-level M-quantile random effects regression has been written in R.<sup>34</sup> Some asymptotic properties of the estimators and their variance parameters for the MQRE models of order  $q$  are discussed in Tzavidis et al.<sup>22</sup>

In order to compare the performance of the estimators from Section 3.2, a Monte-Carlo simulation was conducted. In this simulation, we evaluate the performance of the MQRE-3L 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). The results are reported in supplementary material and they show that MQRE-3L is a good compromise between efficiency under normality and robust properties under contamination. Furthermore, the MQRE-3L performs on a higher level than the MQRE-2L when three-level clustering is present in the data. Finally, for all scenarios we have implemented, 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.



### 3.3 Results

Table 3 shows the estimated three-level MQRE regression coefficients for five quantiles: 0.1, 0.25, 0.5, 0.75 and 0.90. We used the Huber influence function for the MQRE-2L and MQRE-3L with a fixed tuning constant  $c=1.345$ . This value gives reasonably high efficiency under normality and protects against outliers.<sup>21</sup> 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 the change from the baseline 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.

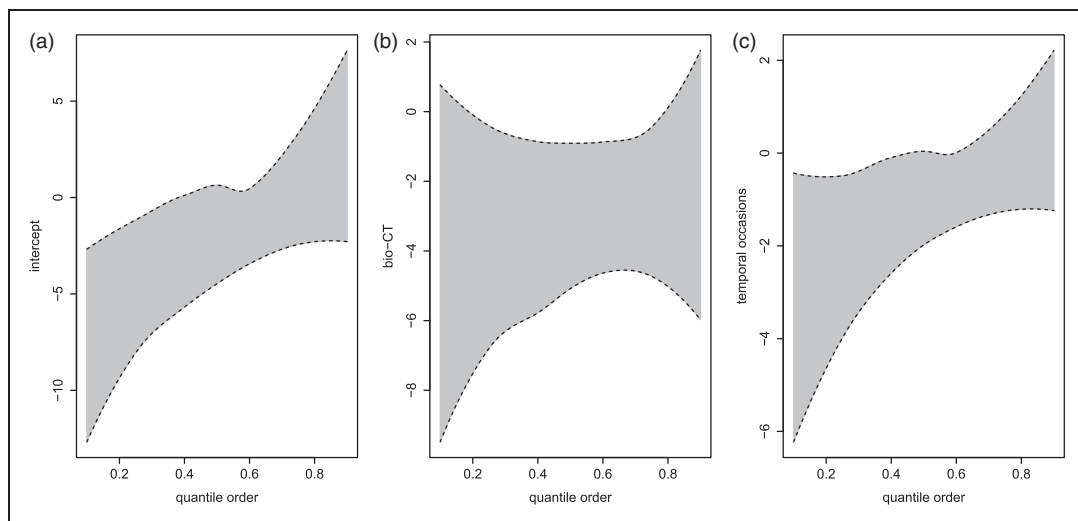
The plots in Figure 6 show the estimated effect of each explanatory variable we included in the model by quantiles. Estimates far from the centre of the distribution usually cannot be evaluated with high precision. To display the sampling variation, a confidence band across the quantiles was constructed 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 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 3 shows that the centre effect ( $\sigma_{centre_q}^2$ ) 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 been 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 raises an issue concerning the heterogeneity induced by

**Table 3.** Parameter estimates and corresponding standard error estimates in parentheses for the data.

	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)
$\sigma_{centre_q}^2$	3.24 (4.08)	6.58 (7.51)	11.97 (7.66)	5.19 (2.59)	0.08 (1.99)
$\sigma_{npz_q}^2$	17.10 (9.28)	35.32 (15.14)	33.03 (7.86)	27.81 (12.05)	20.32 (11.70)
$\sigma_{\epsilon_q}^2$	32.28 (13.52)	36.25 (10.00)	22.99 (4.95)	17.39 (6.39)	14.59 (6.19)

Note: MQRE-3L.



**Figure 6.** MQRE-3L parameter estimates and corresponding 95% confidence interval estimates: (a) intercept, (b) bio-CT treatment and (c) temporal occasion.

potentially 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 MQRE-2L and the LQMM. Table 4 reports the estimated parameters for the MQRE-2L: the magnitude and the sign of the regression coefficients do not change with respect to the MQRE-3L, as well as the values of the estimated variance parameters for the patient and temporal occasion levels. Table 5 shows the estimates of the regression coefficients and variance component for the two-level 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 no differences in the sign of the regression coefficients with respect to 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 with respect to the estimate given by MQRE especially at the tail of the conditional distribution. The MQRE allows the estimation of a two-level and three-level 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-level 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 6 reports the estimated parameter for this

**Table 4.** Parameter estimates and corresponding standard error estimates in parentheses for the data.

	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)
$\sigma_{\eta\beta z_q}^2$	19.10 (9.93)	42.07 (13.72)	45.10 (8.83)	31.99 (12.50)	20.53 (12.33)
$\sigma_{\varepsilon_q}^2$	32.11 (14.78)	36.73 (9.64)	24.10 (4.88)	16.36 (5.40)	14.65 (5.46)

Note: MQRE-2L.

**Table 5.** Parameter estimates and corresponding standard error estimates in parentheses for the data.

	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)
$\sigma_{\eta\beta z_q}^2$	77.21 (—)	61.33 (—)	50.96 (—)	52.52 (—)	48.46 (—)

Note: LQMM (two-level) – random intercepts.

**Table 6.** Parameter estimates and corresponding standard error estimates in parentheses for the data.

	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)
$\sigma_{\eta\beta z_q}^2$	59.94 (—)	55.20 (—)	31.78 (—)	29.46 (—)	49.78 (—)
$\sigma_{time}^2$	10.40 (—)	10.73 (—)	7.68 (—)	7.04 (—)	5.50 (—)

Note: LQMM (two-level) – random slopes.

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-level random intercepts MQRE and two-level LQMM could be appropriate for this application.

#### 4 Discussion and conclusion

The estimates reported in Table 3 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 ( $p\text{-value} = 0.0583$ ).

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 were 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  $-1.38$  ( $SD = 0.60$ ,  $p\text{-value} = 0.0214$ ).

As mentioned above, the MQRE allows for the estimation of a two- and three-level 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 an 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-3L model.

Figure 7 shows the estimated intra-class correlation (ICC) of the MQRE-3L 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-3L 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 is worth to consider further this point.

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.

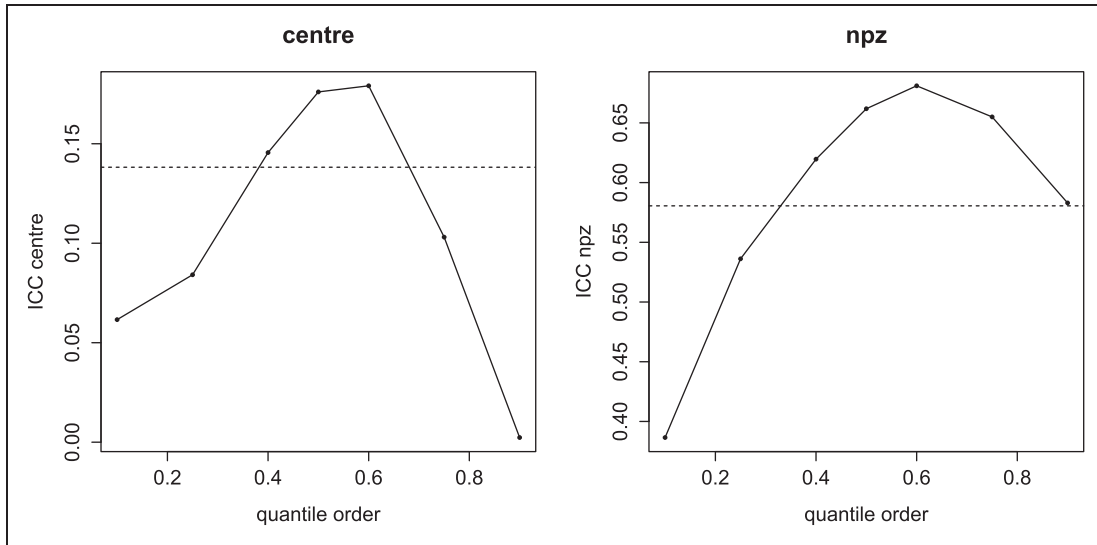


Figure 7. Estimated intra-class correlation of the MQRE-3L at quantiles.

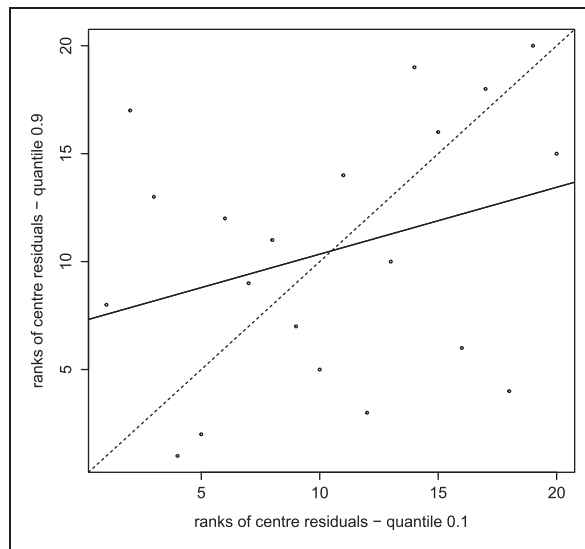
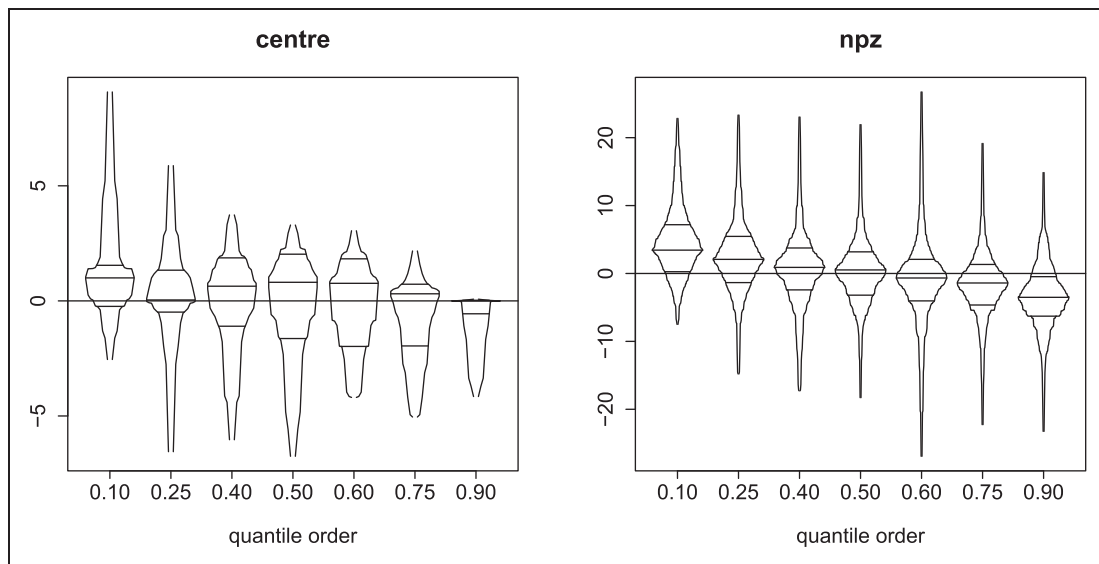


Figure 8. Plot of centre random effect ranks (MQRE-3L for  $q = 0.10, 0.90$ ).

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 it 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.

This seems somehow confirmed by observing that the four centres which lay closer to the first bisector 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



**Figure 9.** Box-centile plots of predicted centre and patient-specific random effects.

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 estimated 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.

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.<sup>35</sup>) 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 and 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.<sup>36</sup> Lipsitz et al.<sup>37</sup> and Yi and He<sup>38</sup> 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.<sup>39</sup> proposed for estimating mean regression. Apart from these two remarkable pieces of work, no other attempts have been made to adjust for non-response in fixed effect quantile models. No papers have dealt with this issue in fixed effect M-quantile modelling. Embedding weights in usual mixed model is a difficult task firstly addressed by Pfeiffermann et al.<sup>40</sup> 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 fixed 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.

### Acknowledgements

Tzavidis gratefully acknowledges the support by grant ES/J001414/1 from the UK Economic and Social Research Council.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by the Progetto di Ricerca di Interesse Nazionale 2012 project ‘Household wealth and youth unemployment: new survey methods to meet current challenges’.

## Supplement Material

Supplementary material is available for this article online.

## References

- Localio AR, Berlin JA and Ten Have TR. Adjustments for Center in Multicenter Studies: an overview. *Ann Int Med* 2001; **135**: 112–123.
- Guthrie L, Oken E, Sterne J, et al. Ongoing monitoring of data clustering in multicenter studies. *BMC Med Res Methodol* 2012; **12**.
- Sakamoto J, Matsuyama Y and Ohashi Y. An analysis of institutional effects in a multicenter cancer clinical trial: is it also plausible from the clinicians’ point of view? *Jpn J Clin Oncol* 1999; **29**: 403–405.
- Yamaguchi T, Ohashi Y and Matsuyama Y. Proportional hazards models with random effects to examine centre effects in multicentre cancer clinical trials. *Stat Meth Med Res* 2002; **11**: 221–236.
- Goldstein H. *Multilevel statistical models*. New York: Wiley, 2003.
- Fairclough DL. *Design and analysis of quality of life studies in clinical trials*. Chapman & Hall/CRC Interdisciplinary Statistics. New York: Taylor & Francis, 2002.
- Koenker R and Bassett G. Regression quantiles. *Econometrica* 1978; **46**: 33–50.
- Koenker R. *Quantile regression*. New York: Cambridge University Press, 2005.
- Carey VJ, Yong FH, Frenkel LM, et al. Growth velocity assessment in paediatric AIDS: smoothing, penalized quantile regression and the definition of growth failure. *Stat Med* 2004; **23**: 509–526.
- Austin PC, Tu JV, Daly PA, et al. The use of quantile regression in health care research: a case study examining gender differences in the timeliness of thrombolytic therapy. *Stat Med* 2005; **24**: 791–816.
- Wei Y, Pere A, Koenker R, et al. Quantile regression methods for reference growth charts. *Stat Med* 2006; **25**: 1369–1382.
- Geraci M and Bottai M. Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics* 2007; **8**: 140–154.
- Bottai M, Cai B and McKeown RE. Logistic quantile regression for bounded outcomes. *Stat Med* 2010; **29**: 309–317.
- Li Y, Graubard BI and Korn EL. Application of nonparametric quantile regression to body mass index percentile curves from survey data. *Stat Med* 2010; **29**: 558–572.
- Nicholson W, Setse R, Hill-Briggs F, et al. Depressive symptoms and health-related quality of life in early pregnancy. *Obstet Gynecol* 2006; **107**: 798–806.
- Pourhoseingholi MA, Safaee A, Moghimi-Dehkordi B, et al. Quality of life in breast cancer patients – a quantile regression analysis. *Asian Pacific J Cancer Prevent* 2008; **9**: 487–490.
- Broccoli S, Cavrini G and Zoli M. A quantile regression approach to the analysis of the quality of life determinants in the elderly. *Statistica* 2005; **65**: 419–436.
- Geraci M and Bottai M. Linear quantile mixed models. *Stat Comput* 2014; **24**: 461–479.
- Breckling J and Chambers R. M-quantiles. *Biometrika* 1988; **75**: 761–771.
- Newey WK and Powell JL. Asymmetric least squares estimation and testing. *Econometrica* 1987; **55**: 819–847.
- Huber P. *Robust statistics*. New York: Wiley, 1981.
- Tzavidis N, Salvati N, Schmid T, et al. Longitudinal analysis of the Strengths and Difficulties Questionnaire scores of the Millennium Cohort Study children in England using M-quantile random effects regression. *J R Stat Soc: Ser A* 2016; **179**: 427–452.
- Jones MC. Expectiles and M-quantiles are quantiles. *Stat Prob Lett* 1994; **20**: 149–153.
- Ridolfi R, Chiarion-Sileni V, Guida M, et al. 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. *J Clin Oncol* 2002; **20**: 1600–1607.
- Chiarion-Sileni V, Del Bianco P, De Salvo GL, et al. Quality of life evaluation in a randomised trial of chemotherapy versus bio-chemotherapy in advanced melanoma patients. *Eur J Cancer* 2003; **39**: 1577–1585.
- Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. *J Am Stat Assoc* 1977; **72**: 320–338.
- Richardson AM and Welsh AH. Robust restricted maximum likelihood in mixed linear models. *Biometrics* 1995; **51**: 1429–1439.
- Huggins RM. A robust approach to the analysis of repeated measures. *Biometrics* 1993; **49**: 715–720.
- Richardson AM. Bounded influence estimation in the mixed linear model. *J Am Stat Assoc* 1997; **92**: 154–161.

30. Huggins RM and Loesch DZ. On the analysis of mixed longitudinal growth data. *Biometrics* 1998; **54**: 583–595.
31. Sinha SK and Rao JNK. Robust small area estimation. *Can J Stat* 2009; **37**: 381–399.
32. Anderson TW. Asymptotically efficient estimation of covariance matrices with linear structure. *Ann Stat* 1973; **1**: 135–141.
33. Staudenmayer J, Lake E and Wand M. Robustness for general design mixed models using the t-distribution. *Stat Model* 2009; **9**: 235–255.
34. R Development Core Team. R: a language and environment for statistical computing, Vienna, www.R-project.org (2011, accessed 19 February 2016).
35. le Cessie S, de Vries EGE, Buijs C, et al. Analyzing longitudinal data with patients in different disease states during follow-up and death as final state. *Stat Med* 2009; **28**: 3829–3843.
36. Little RJA and Rubin DB. *Statistical analysis with missing data*. New York, NY, USA: John Wiley & Sons, Inc., 2002.
37. Lipsitz SR, Fitzmaurice GM, Molenberghs G, et al. Quantile regression methods for longitudinal data with drop-outs: application to CD4 cell counts of patients infected with the human immunodeficiency virus. *J R Stat Soc: Ser C* 1997; **46**: 463–476.
38. Yi GY and He W. Median regression models for longitudinal data with dropouts. *Biometrics* 2009; **65**: 618–625.
39. Robins JM, Rotnitzky A and Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc* 1995; **90**: 106–121.
40. Pfeffermann D, Skinner CJ, Holmes DJ, et al. Weighting for unequal selection probabilities in multilevel models. *J R Stat Soc: Ser B* 1998; **60**: 23–40.