Qual Life Res
DOI 10.1007/s11136-014-0808-3

Improved statistical analysis of pre- and post-treatment patient-reported outcome measures (PROMs): the applicability of piecewise linear regression splines

M. E. Greene · O. Rolfson · G. Garellick · M. Gordon · S. Nemes

Accepted: 15 September 2014

Abstract
Purpose Patient-reported health-related quality-of-life (HRQoL) measures such as the EuroQol 5 dimension (EQ-5D) index are commonplace when assessing healthcare providers or efficiency of medical techniques. HRQoL measures are generally bounded, and the magnitude of possible improvement depends on the pre-treatment HRQoL value. This paper aimed to assess and illustrated the possibility of modelling the relationship between pre- and post-treatment HRQoL measures with piecewise linear splines.
Methods The method was illustrated using a longitudinal dataset of 36,625 patients with one EQ-5D index before and one a year after total hip arthroplasty. We considered four models: intercept only model, single line regression, and segmented regression with 1 and 2 change points. The post-operative EQ-5D index served as the outcome, while the preoperative EQ-5D index was the predictor.
Results We found that a two-line regression best described the data with the lines meeting at 0.159 on the preoperative EQ-5D index scale. In the low preoperative group (with an initial preoperative index from −0.594 to 0.159), the predicted post-operative scores ranged from 0.368 to 0.765, with post-operative scores increasing 0.528 points for each unit in the preoperative score. In the high preoperative group (initial range from 0.159 to 1), the predicted post-operative scores ranged from 0.765 to 0.855, increasing 0.106 points for each unit in the preoperative score.
Conclusions Piecewise linear regression is a straightforward approach to analyse baseline and follow-up HRQoL measures such as the EQ-5D index. It can provide a reasonable approximation of the shape of the underlying relationship where the threshold and slopes prove informative and meaningful.

Keywords EQ-5D · Segmented regression · Change point · Total hip arthroplasty

Electronic supplementary material The online version of this article (doi:10.1007/s11136-014-0808-3) contains supplementary material, which is available to authorized users.

M. E. Greene · O. Rolfson · G. Garellick · M. Gordon · S. Nemes (✉)
Swedish Hip Arthroplasty Register, Centre of Registers Västra Götaland, Medicinargatan 18G, 413 45 Gothenburg, Sweden
e-mail: szilard.nemes@registercentrum.se

M. E. Greene · O. Rolfson · G. Garellick · S. Nemes
Department of Orthopedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, 41345 Gothenburg, Sweden

M. E. Greene
Harris Orthopaedic Laboratory, Massachusetts General Hospital, Boston, MA, USA

M. Gordon
Department of Clinical Sciences at Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden

Introduction

Regression models with change points have a long history [1–4], and with the availability of accessible computing power, experienced a resurgence that led to numerous theoretical/statistical advances making the applicability of the piecewise linear regression splines easy and straightforward [5–11]. Piecewise linear regression splines found their applicability in fields ranging from epidemiology to genetics [12, 13].

Published online: 25 September 2014

Patient-reported health-related quality-of-life (HRQoL) measures have the potential to act as drivers for better organization and improvement of health care [14, 15].

Despite widespread usage of self-assessed HRQoL measures, there is no consensus in the scientific community on how to best obtain clinically and economically meaningful effect measures from them. The premise is that an intervention will likely change the self-assessed HRQoL of the patients. However, the degree of change may depend on the HRQoL prior to the intervention. Patients with poor preoperative HRQoL might experience substantial improvement, while patients with higher preoperative self-assessed HRQoL might show only a minute increase or even stagnation.

A common use of the generic EuroQol 5 dimension HRQoL questionnaire (EQ-5D) [16] is to compare populations over time and to gauge the cost effectiveness of medical treatments [17]. When measuring HRQoL with the EQ-5D index, this effect may be exaggerated by the floor and ceiling of the EQ-5D index. According to the British time-trade-off (TTO) value set, the index cannot go beyond −0.594 and 1.000 [18]. Researchers who aim to understand the relationship between HRQoL measurements prior to and after clinical interventions have to assure that the predicted scores are in the appropriate range.

In this paper, we explore the use of piecewise linear regression splines for the analysis of HRQoL measures, with specific interest in EQ-5D index profiles. Regression models with change points assume that the structural relationship between the outcome and the predictor will differ in various patient groups within the dataset [19, 20]. Application of piecewise linear regression splines has the advantage of fitting separate regression equations to different subsets of the data. Providing a vivid description of changes in HRQoL measures will aid clinicians and researchers trying understand the influence of an intervention on patient-reported HRQoL. We illustrate the applicability of our proposed method with pre- and post-operative EQ-5D indices of 36,625 Swedish patients treated with total hip arthroplasty (THA).

Methods

Data collection and measures

In Sweden, patients undergoing THA are routinely monitored by a patient-reported outcome measures (PROMs) programme run by the Swedish Hip Arthroplasty Register. Patients are contacted preoperatively by their clinic and 1 year post-operatively by mail to complete the surveys. Response rates are 86 % preoperatively, and 90 % of the patients who return the preoperative questionnaire participate at the one-year follow-up as well [21]. In order to achieve high response rates, the questionnaire is kept short comprising a visual analogue scale for hip pain, the Charnley classifier for musculoskeletal comorbidity [22, 23], and the HRQoL measure EuroQol 5 dimension HRQoL questionnaire (EuroQolGroup 1990). The EuroQol 5 dimension HRQoL questionnaire consists of five dimensions measuring mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. With increasing severity, there are three answer options for each dimension: (1) no problems, (2) moderate problems, and (3) extreme problems. This yields 243 unique health states which can be used to calculate a single, weighted health utility index where 1 is perfect HRQoL, 0 represents death, and scores <0 are health states worse than death. Patients who respond that they have 'extreme problems' in all five dimensions receive the lowest possible score of −0.594. Different valuation sets are used to adjust for cultural response pattern differences. Lacking a specific Swedish value set, the Swedish Hip Arthroplasty Register uses the British TTO value set to calculate the EQ-5D index in their population. The EuroQol 5 dimension HRQoL questionnaire is commonly utilized in the clinical setting as a general health measure because it is short, easy for patients to interpret, and has been validated in over 150 languages and is frequently used to conduct cost effectiveness analyses. There are extensive published works confirming the validity and reliability of this tool in clinical and research settings [24–26].

Data from 36,625 primary THA patients with a diagnosis of primary osteoarthritis performed between 2002 and 2011 with complete preoperative and one-year post-operative PROMs were selected from the Swedish Hip Arthroplasty Register. Each patient contributed two EQ-5D indices one before and one a year after total hip arthroplasty.

Primary osteoarthritis is the main reason for THA surgeries, accounting for around 82 % of cases. Facture patients who account for around 10 % of the cases were excluded as they lack preoperative PROM values. We included patients with the most common surgical techniques. If a patient had arthroplasty on both hips during the study period, only the first operation with complete PROMs was selected. The exclusions were made to reduce residual confounding, and these data were selected for illustrative purposes. Ethical review approval was obtained from the Central Ethical Review Board in Gothenburg, Sweden (decision 293-13).

Model formulation

The effect of preoperative EQ-5D index values on the post-operative EQ-5D index values were explored with four regression equations. We considered four models: intercept only null model, single line regression, and segmented regressions with 1 and 2 change points. The post-operative EQ-5D index served as the outcome for the regression modelling, while the preoperative EQ-5D index was the predictor.

Generally, there are two possible situations with which the relationship between the outcome and predictors can be best described with segmented models; a continuous piecewise model with regression lines having different slopes connected at unknown change points or a discontinuous regression line that jumps at change point values.

In the following, we restrict our attention to the one change point regression equation, where the outcome is modelled as

$$y_i = \begin{cases} \beta_0^{(1)} + \beta_1^{(1)} x_i + \varepsilon_{i1} & \text{if} \quad x_i < r \\ \beta_0^{(2)} + \beta_1^{(2)} x_i + \varepsilon_{i2} & \text{if} \quad x_i > r \end{cases} \tag{1}$$

where r is the unknown change point, and $\beta_0^{(1)}$, $\beta_1^{(1)}$ and $\beta_0^{(2)}$, $\beta_1^{(2)}$ are the intercept and slope for equations on the left and right side of the change point, respectively. Each of the model parameters (change point, intercept, and slope) was estimated from the data.

If we assume that the relationship between the outcome and predictor is continuous, then the following restriction is imposed

$$\beta_0^{(1)} + \beta_1^{(1)} r = \beta_0^{(2)} + \beta_1^{(2)} r \tag{2}$$

If the continuous model is assumed, then Eq. 1 can be written as

$$y_i = \beta_0 + \beta_1^* x_i + \beta_r^* (x_i - r) I(x_i < r) + \varepsilon_i \tag{3}$$

where $\beta_0 = \beta_0^{(1)}$, $\beta_1^* = \beta_1^{(1)}$, $\beta_r^* = \beta_1^{(2)} - \beta_1^{(1)}$, and $\beta_0^{(2)} = \beta_0^{(1)} + r(\beta_1^{(1)} - \beta_1^{(2)})$. In this case, $I(x_i < r)$ is an indicator function which takes a value of 1 if the condition is met; otherwise, it is 0. The superscript asterisk (*) denotes a model parameter that is defined as the difference between the two slopes. For the matrix notation, we refer the reader to the Supplementary materials.

Parameter estimation and statistical inference require the definition of a loss function, either likelihood or least squares. Conditional on the change points the density function for the outcome y_i will be

$$f(\cdot|\boldsymbol{\beta}_r) = f_1(\cdot|\boldsymbol{\beta}_r)^{I(x<r)} f_2(\cdot|\boldsymbol{\beta}_r)^{I(x>r)} \tag{4}$$

and the likelihood function will be

$$\ell(\boldsymbol{\beta}_r) = \sum_{i=1}^{n} \{ I(x<r) \ln f_1(\cdot|\boldsymbol{\beta}_r) + I(x>r) \ln f_2(\cdot|\boldsymbol{\beta}_r) \}. \tag{5}$$

The density functions are assumed to be Gaussian. Parameter estimates are given by $\boldsymbol{\beta}_r$ that maximize the log-likelihood. There is no single best method for optimization; the choices are influenced by the assumed structural form whether continuous or not [19]. We chose the linearization proposed by Muggeo [27] which does not assume constant variance across regions. The Muggeo's linearization provides simultaneous parameter estimation and statistical inference, and its estimates are more accurate than the other possible alternatives [19]. It should be noted that the EQ-5D index data violate the normality assumption, thus in small samples, extra caution is warranted. Computer intensive methods such as bootstrapping might be useful to consider with this kind of data.

Model selection

Model selection was based on the Bayesian Information Criterion that penalizes additional parameters with a factor of $\ln(n)$, where n is the sample size. Once the working model was selected, the preoperative index values were used to calculate the range of predicted post-operative index values as a means to describe how well the model fit the data.

Model validation

To ascertain whether the model would accurately predict patient outcomes, we utilized a validation procedure based on the 0.632 bootstrap [28, 29]. A total of 1,000 bootstrap resamples were drawn from the original dataset. The 0.632 bootstrap assumes a resampling probability of 1–1/n where n is the sample size. The models are refitted on the bootstrapped data with the original coefficient estimates. Differences in R^2 values between the original and resampled data denote over-optimism in the estimation. If this over-optimism is substantial, we have an indication of over-fitting when the fit of a complex model cannot be adequately replicated in subsequent samples or resamples. For additional details, we refer the reader to the book by Harrell [30].

Adjustment

To see if the estimated change point withstands adjustment for confounders, we adjusted the model for two background variables (age and gender), a clinical variable (surgical approach) and the preoperative Charnley classifier for musculoskeletal comorbidities. Preoperative EQ-5D index and age were modelled as piecewise linear splines.

Software

Calculations were run using the R computing environment R 3.0.1, R Core Team 2013 [31], and we used the 'segmented' package for change point estimation [27, 32].

Table 1 Parameters of the three competing models considered for modelling the changes in post-operative EQ-5D index as a function of the preoperative index values

		Null	Linear	2-Segment	
				First segment	Second segment
Intercept	β (±1 SE)	0.777 (0.001)	0.701 (0.002)	0.682 (0.002)	0.749 (0.002)
	95 % CI	0.774; 0.779	0.697; 0.705	0.677; 0.686	0.744; 0.753
Slope	β (±1 SE)	–	0.181 (0.003)	0.528 (0.02)	0.106 (0.008)
	95 % CI	–	0.174; 0.189	0.486; 0.568	0.089; 0.129
R^2		–	5.67 %	6.58 %	
BIC		−1,434.02	−3,565.15	−3,901.07	

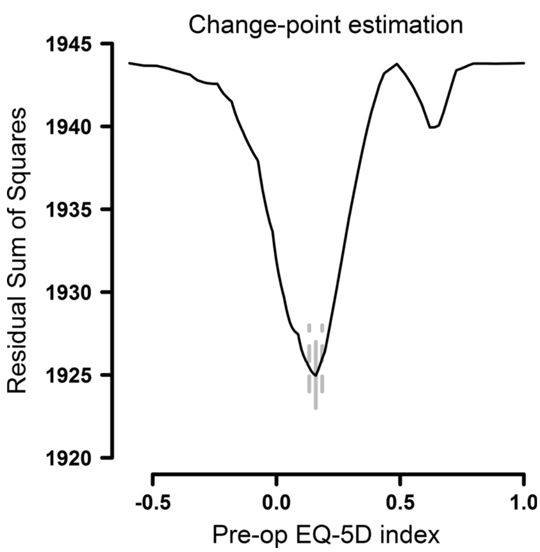

Fig. 1 Change point estimation by minimizing the residual sum of squares with the estimated change point of a two-segment piecewise linear regression spline and the associated 95 % confidence intervals

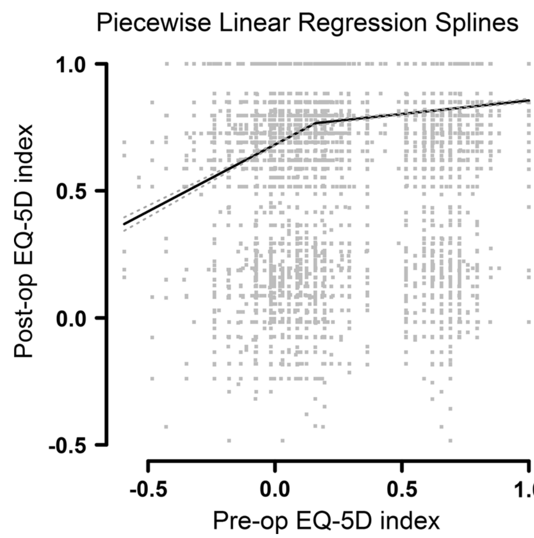

Fig. 2 Piecewise linear regression spline modelling of pre- and post-operative EQ-5D indices with the associated 95 % confidence intervals. The relationship is best described by a two-segment regression equation joined at a preoperative EQ-5D index change point of 0.159

Results

Of the 36,625 patients studied, 30,807 (84.1 %) showed improved post-operative EQ-5D indices, 2,557 (6.9 %) did not experience any change, while 3,261 (9.0 %) deteriorated in HRQoL. Regression modelling and model selection based on BIC suggested that the two-line regression provided the best fit for the data (Table 1). The estimated change point was at EQ-5D index value of 0.159 (95 % CI 0.135–0.182) as illustrates in Fig. 1. We did not observe any support for a three-line regression and estimation of two change points failed, and the algorithm did not produce parameter estimates.

For the purpose of reporting, individuals with a preoperative EQ-5D index of 0.159 or less were classified as having low preoperative HRQoL (n = 13,157 (35.9 %)) and those over 0.159 as having high preoperative HRQoL (n = 23,468 (64.1 %)).

In the low preoperative group (with preoperative EQ-5D index values from −0.594 to 0.159), the predicted post-operative EQ-5D index values ranged from 0.368 to 0.765, with post-operative EQ-5D index values increasing 0.528 (95 % CI 0.480–0.576) points for each unit in the preoperative EQ-5D index values. In the high preoperative group (initial range from 0.159 to 1.000), the predicted post-operative EQ-5D index values ranged from 0.765 to 0.855, increasing 0.106 (95 % CI 0.091–0.121) points for each unit in the preoperative EQ-5D index values (Fig. 2). Table 2 summarizes the predicted values for the post-operative EQ-5D index values based on preoperative EQ-5D index values with simple linear regression and piecewise linear regression splines. The increment was set at 0.1, and the expected increase with 0.1 units of change in preoperative EQ-5D index by the simple linear regression was 0.018. The expected increase with 0.1 units of change in preoperative EQ-5D index by piecewise linear

Table 2 Predicted post-operative EQ-5D index values and associated 95 % confidence intervals based on simple linear regression and by two-segment piecewise linear regression splines

EQ-5D index	Linear		Piecewise	
	Predicted	95 % CI	Predicted	95 % CI
−0.5	0.611	0.604–0.618	0.416	0.390–0.443
−0.4	0.629	0.622–0.636	0.469	0.448–0.491
−0.3	0.647	0.641–0.653	0.523	0.506–0.539
−0.2	0.665	0.660–0.671	0.576	0.564–0.588
−0.1	0.684	0.679–0.688	0.629	0.621–0.637
0.0	0.702	0.698–0.706	0.682	0.678–0.687
0.1	0.720	0.717–0.723	0.735	0.730–0.741
0.2	0.738	0.735–0.741	0.770	0.765–0.775
0.3	0.756	0.754–0.759	0.781	0.776–0.785
0.4	0.774	0.772–0.777	0.791	0.788–0.795
0.5	0.793	0.790–0.795	0.802	0.799–0.805
0.6	0.811	0.808–0.813	0.813	0.810–0.816
0.7	0.829	0.826–0.832	0.824	0.821–0.827
0.8	0.847	0.843–0.851	0.834	0.830–0.839
0.9	0.865	0.861–0.870	0.845	0.840–0.850
1.0	0.883	0.878–0.888	0.856	0.850–0.862

regression splines was 0.052 for values under 0.159 and 0.01 for values 0.159 and higher.

The predictive power of the model was 6.59 %. The model validation procedure did not identify problems with over-fitting. The mean predictive power of the model in the training set was 6.60 %, while in the validation set 6.55 %, indicating an over-estimation of 0.75 %.

Adjustment for important confounders did not induced substantial change in the estimated change point and resulted in minor changes in the slope coefficients. BIC indicated that a model with one change point for the EQ-5D index at 0.158 (95 % CI 0.133–0.183) and one change point for age at 68.98 years (95 % CI 66.73–71.23) offered the best fit (Table 3). The model's predictive power was 12.67 %. This was largely due to the inclusion of Charnley classification (partial $R^2 = 4.69$ %), despite significance age, gender, and surgical approach had negligible influence.

Table 3 Parameters of the multivariate model with post-operative EQ-5D index as the outcome and preoperative EQ-5D index as the exposure

	r	95 % CI	β (±1 SE)	95 % CI	Partial R^2
Intercept	–	–	0.790 (0.02)	0.746; 0.833	–
Gender (women)	–	–	−0.023 (0.002)	−0.028; −0.018	0.23 %
Surgical approach	–	–	−0.028 (0.002)	−0.033; −0.023	0.34 %
Charnley class					
A					4.69 %
B	–	–	−0.058 (0.003)	−0.066; −0.051	
C	–	–	−0.111 (0.002)	−0.116; −0.106	
Age					
First	68.98	66.73; 71.23	0.00001 (0.0003)	−0.0006; 0.0007	0.95 %
Second	–		−0.003 (0.0002)	−0.004; −0.0029	
EQ-5D index					
First	0.158	0.133; 0.183	0.471 (0.024)	0.422; 0.518	4.81 %
Second	–		0.081 (0.006)	0.068; 0.093	

The model is adjusted for gender (reference male), surgical approach (posterior as reference), Charnley classification, and age. Preoperative EQ-5D index and age are modelled as piecewise linear regression splines. The table presents estimates (±1 standard error) and associated 95 % confidence intervals for the change point r, the intercept, and slope β. The partial R^2 represents the loss in predictive power if a specific variable is removed from the model

Discussion

In this paper, we demonstrate that piecewise linear regression splines provide a better model fit and better predictive power relative to simple linear regression when HRQoL measures such as the EQ-5D index in a nationwide cohort of THA patients is analysed. The selected model accounts for the fact that the same post-operative effect is not expected through the entire range of the preoperative EQ-5D indices. An added benefit of the different slopes is that predicted post-operative scores based on the preoperative scores are more likely to be in the appropriate range. Thus, piecewise linear regression splines are more likely to conform to the floor and ceiling effect that simple linear regression cannot, although they do not have the ability to explicitly consider these limits. Alternatives like nonlinear least squares regression consider not only explicitly the existence of the limits but also the ability to estimate their value from the data [33, 34]. However, these models were derived for modelling growth rates and disease progression, so the relationship that they assume between the predictor and outcome might not be proper for HRQoL studies. Censored or truncated regression models may also consider the existence of floor and ceiling values [35, 36] which can be set by the researcher, yet they assume a constant effect throughout the whole range of outcome scores from the predictor. Regression splines such as restricted cubic splines do not necessarily consider the existence of floor and ceiling values, but they will likely conform to the data well. The main drawback of such splines is the lack of presentable and interpretable effect measures. Nonlinear splines are routinely reported in graphical form, thus hampering between-study comparisons and or meta-analyses. An added challenge with nonlinear splines is the lack of generally accepted and proven statistical routines for statistical inference. Classic tests such as the ANOVA are hard to use and interpret with splines, and Information Theoretic Criteria (AIC or BIC) might not guard satisfactorily against artefacts.

Piecewise linear regression splines have attractive mathematical/statistical characteristics allowing researchers to interpret the change point as a distinguishing characteristic between patient groups preoperatively. This change point might be abrupt, nonlinear splines might offer a smoother transition point. However, only piecewise linear regression splines can exactly estimate the location of such points when the structural relationship between predictor and outcome changes. In this paper, we focused on the relationship between preoperative and post-operative values of the same measure. As we illustrated with age, this method also fits well with other exposures. Albeit a very weak relationship between age and post-operative EQ-5D index, we were able to identify that patients aged over 69 years benefited less from THA than younger patients. Considering further measures such as BMI and applying them to large cohorts might lead to the determination of important thresholds that have the potential to aid in clinical decision-making.

Analysis of change scores (post-operative value minus the preoperative value) with preoperative values as a predictor still prevails in the literature; however, they can lead to inaccurate results because in this method, the change score is negatively correlated with the preoperative data [37, 38].

Piecewise linear regression splines can be applied as a standalone statistical analysis, and interpretation of the regression coefficients can offer essential insights. In this analysis, we show that the slopes of the regression lines represent the influence of preoperative HRQoL on post-operative recovery and general wellbeing. The coefficient of determination (R^2) tells how much of the variation in the post-operative EQ-5D index was due to the preoperative HRQoL. In our case, around 6.5 % of the variation in post-operative EQ-5D index was explained by preoperative HRQoL. This suggests that factors other than preoperative HRQoL also need to be considered. Adjusting the model for age, gender, surgical approach, and preoperative musculoskeletal comorbidity (Charnley classification) increased the coefficient of determination, mostly due to the addition of Charnley classification.

Often, the purpose of studying HRQoL is to compare alternative treatments or approaches. In most cases, patients cannot be randomly allocated to different treatments types as they would be in a randomized clinical trial. Thus, to be able to compare alternative treatments for the same condition, we need to conduct proper risk adjustments [39] and the better we describe the relationship between the outcome and predictors, the less likely we are to induce bias in the parameter of interest [40]. Piecewise linear regression splines offer a simple and easy to interpret approximation of this unknown relationship. Alternatives, such as nonparametric regression or generalized additive models, might provide a better mathematical fit. However, we believe that simplicity and ease of interpretation of the piecewise linear regression outweigh the gains in the goodness of fit of the aforementioned methods. It is intuitive that patients who report poor preoperative HRQoL (those with low preoperative EQ-5D index scores) have more room for improvement on the EQ-5D index scale after treatment. Contrastingly, patients who report 'no problems' in many of the five dimensions before treatment have little to no capacity for improvement when measured by the EQ-5D index. Thus, even if the perceived improvement in HRQoL is substantial and the operation was life changing, this improvement will induce little to no change in the EQ-5D index. By dividing the patients into two (or more) groups, we can have a better understanding of the benefits of a given treatment.

Conclusions

Piecewise linear regression splines are a useful and practical approach to analysing and reporting HRQoL measures such as the EQ-5D index. The main appeal of segmented models lies in the easy interpretability of the influence of the preoperative HRQoL on post-operative scores in different patient subgroups. Piecewise linear regression splines may provide a reasonable approximation of the shape of the underlying pre- and post-operative HRQoL relationship where the threshold and slopes prove informative and meaningful [32].

Acknowledgments This study is an analysis of the databases in the Swedish Hip Arthroplasty Register, which is funded by the Swedish National Board of Health and Welfare, the Swedish Association of Local Authorities and Regions and the County Council of the Western Region of Sweden. Research grants from Dr Felix Neuberghs Foundation have contributed financially to this study. The publication of the results is not contingent on the sponsor's approval.

References

1. Quandt, R. E. (1958). The estimation of the parameters of a linear-regression system obeying 2 separate regimes. *Journal of the American Statistical Association, 53*(284), 873–880.
2. Quandt, R. E. (1960). Tests of the hypothesis that a linear-regression system obeys 2 separate regimes. *Journal of the American Statistical Association, 55*(290), 324–330.

3. Hinkley, D. V. (1969). Inference about intersection in 2-phase regression. *Biometrika, 56*(3), 495–504.
4. Hinkley, D. V. (1971). Inference in 2-phase regression. *Journal of the American Statistical Association, 66*(336), 736–743.
5. Kim, H. J., Fay, M. P., Feuer, E. J., & Midthune, D. N. (2000). Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine, 19*(3), 335–351.
6. Kim, H. J., Fay, M. P., Yu, B. B., Barrett, M. J., & Feuer, E. J. (2004). Comparability of segmented line regression models. *Biometrics, 60*(4), 1005–1014.
7. Kim, H. J., Yu, B., & Feuer, E. J. (2008). Inference in segmented line regression: A simulation study. *Journal of Statistical Computation and Simulation, 78*(11), 1087–1103.
8. Kim, J., & Kim, H. J. (2008). Asymptotic results in segmented multiple regression. *Journal of Multivariate Analysis, 99*(9), 2016–2038.
9. Liu, Z. H., & Qian, L. F. (2010). Changepoint estimation in a segmented linear regression via empirical likelihood. *Communications in Statistics-Simulation and Computation, 39*(1), 85–100.
10. He, H. P., & Severini, T. A. (2010). Asymptotic properties of maximum likelihood estimators in models with multiple change points. *Bernoulli, 16*(3), 759–779.
11. Julious, S. A. (2001). Inference and estimation in a changepoint regression problem. *Journal of the Royal Statistical Society Series D-the Statistician, 50*, 51–61.
12. Pastor, R., & Guallar, E. (1998). Use of two-segmented logistic regression to estimate change-points in epidemiologic studies. *American Journal of Epidemiology, 148*(7), 631–642.
13. Nemes, S., Parris, T. Z., Danielsson, A., Kannius-Janson, M., Jonasson, J. M., Steineck, G., et al. (2012). Segmented regression, a versatile tool to analyze mRNA levels in relation to DNA copy number aberrations. *Genes, Chromosomes and Cancer, 51*(1), 77–82.
14. Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ, 346*, f167.
15. Rolfson, O., Rothwell, A., Sedrakyan, A., Chenok, K. E., Bohm, E., Bozic, K. J., et al. (2011). Use of patient-reported outcomes in the context of different levels of data. *The Journal of Bone and Joint Surgery, 93*(Supplement_3), 66–71.
16. Szende, A., Oppe, M., & Devlin, N. (2007). *EQ-5D value sets: Inventory, comparative review and user guide*. EuroQol Group Monographs (Vol. 2). Dordrecht: Springer.
17. Smith, P. C., & Street, A. D. (2013). On the uses of routine patient-reported outcome data. *Health Economics, 22*(2), 119–131.
18. Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35*(11), 1095–1108.
19. Chen, C. W. S., Chan, J. S. K., Gerlach, R., & Hsieh, W. Y. L. (2011). A comparison of estimators for regression models with change points. *Statistics and Computing, 21*(3), 395–414.
20. Chen, C. W. S., Chan, J. S. K., So, M. K. P., & Lee, K. K. M. (2011). Classification in segmented regression problems. *Computational Statistics and Data Analysis, 55*(7), 2276–2287.
21. Rolfson, O., Karrholm, J., Dahlberg, L. E., & Garellick, G. (2011). Patient-reported outcomes in the Swedish Hip Arthroplasty Register. Results of a nationwide prospective observational study. *Journal of Bone and Joint Surgery-British, 93B*(7), 867–875.
22. Callaghan, J., Dysart, S., Savory, C., & Hopkinson, W. (1990). Assessing the results of hip replacement. A comparison of five different rating systems. *Journal of Bone and Joint Surgery, British, 72-B*(6), 1008–1009.
23. Charnley, J. (1979). *Low friction arthroplasty of the hip*. Theory and practice. Berlin: Springer.
24. Rabin, R., & de Charro, F. (2001). EQ-5D: A measure of health status from the EuroQol group. *Annals of Medicine, 33*(5), 337–343.
25. Dyer, M. T. D., Goldsmith, K. A., Sharples, L. S., & Buxton, M. J. (2010). A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health and Quality of Life Outcomes, 8*, 12.
26. Pickard, A. S., Wilke, C. T., Lin, H. W., & Lloyd, A. (2007). Health utilities using the EQ-5D in studies of cancer. *Pharmacoeconomics, 25*(5), 365–384.
27. Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine, 22*(19), 3055–3071.
28. Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association, 78*(382), 316–331.
29. Efron, B., & Tibshirani, R. (1997). Improvements on cross-validation: The .632+ bootstrap method. *Journal of the American Statistical Association, 92*(438), 548–560.
30. Harrell, F. (2001). Regression modeling strategies. With applications to linear models, logistic regression and survival analysis. Springer.
31. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (http://www.R-project.org/).
32. Muggeo, V. M. R. (2008). 'Segmented': An R package to fit regression models with broken-line relationships. *RNews, 8*, 20–25.
33. Nemes, S., Gordon, M., Rogmark, C., & Rolfson, O. (2014). Projections of total hip replacement in Sweden from 2013 to 2030. *Acta Orthopaedica, 85*(3), 238–243.
34. Paine, C. E. T., Marthews, T. R., Vogt, D. R., Purves, D., Rees, M., Hector, A., et al. (2012). How to fit nonlinear plant growth models and calculate growth rates: An update for ecologists. *Methods in Ecology and Evolution, 3*(2), 245–256.
35. Oman, S. D. (1983). Regression estimation for a bounded response over a bounded region. *Technometrics, 25*(3), 251–261.
36. Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica, 26*(1), 24–36.
37. Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20*, 93–114.
38. Vickers, A. J., & Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *BMJ, 323*(7321), 1123–1124.
39. Smith, P. C., & Street, A. D. (2013). On the uses of routine patient-reported health outcome data. *Health Economics, 22*(2), 119–131.
40. Steineck, G., Hunt, H., & Adolfsson, J. (2006). A hierarchical step-model for causation of bias-evaluating cancer treatment with epidemiological methods. *Acta Oncologica, 45*(4), 421–429.