

Heigenhauser, Ludwig; Confavreux, C.; Daumer, Martin; Ebers, C. G.; Kappos, Ludwig

Working Paper

Treating systematic errors in multiple sclerosis data

Discussion Paper, No. 435

Provided in Cooperation with:

Collaborative Research Center (SFB) 386: Statistical Analysis of discrete structures - Applications in Biometrics and Econometrics, University of Munich (LMU)

Suggested Citation: Heigenhauser, Ludwig; Confavreux, C.; Daumer, Martin; Ebers, C. G.; Kappos, Ludwig (2005) : Treating systematic errors in multiple sclerosis data, Discussion Paper, No. 435, Ludwig-Maximilians-Universität München, Sonderforschungsbereich 386 - Statistische Analyse diskreter Strukturen, München, <https://doi.org/10.5282/ubm/epub.1804>

This Version is available at:

<https://hdl.handle.net/10419/31129>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Treating Systematic Errors in Multiple Sclerosis Data

L Heigenhauser¹, C Confavreux², M Daumer³, G Ebers⁴,
L Kappos⁵, C Lederer¹, A Neiß¹, C Polman⁶, S Vukusic²,
B Hellriegel⁷

April 21, 2005

¹Sylvia Lawry Centre for Multiple Sclerosis Research, Munich, Germany

²Hôpital Neurologique, Lyon, France

³Correspondence: Martin Daumer, Sylvia Lawry Centre for Multiple Sclerosis Research, Hohenlindenerstr. 1, 81677 München, Germany, email daumer@slcmsr.org

⁴Department of Clinical Neurology, Oxford University, UK

⁵Neurologisch-Neurochirurgische Poliklinik, Kantonsspital, Universität Basel, Switzerland

⁶Department of Neurology, Vrije University, Amsterdam, Netherlands

⁷Institut für Medizinische Statistik und Epidemiologie, Technical University Munich, Germany

Abstract

Multiple sclerosis (MS) is characterized by high variability between patients and, more importantly here, within an individual over time. This makes categorization and prognosis difficult. Moreover, it is unclear to what degree this intra-individual variation reflects the long-term course of irreversible disability and what is attributable to short-term processes such as relapses, to interrater variability and to measurement error. Any investigation and prediction of the medium or long term evolution of irreversible disability in individual patients is therefore confronted with the problem of systematic error in addition to random fluctuations. The approach described in this article aims to assist in detecting relapses in disease curves and in identifying the underlying disease course. To this end neurological knowledge was transformed into simple rules which were then implemented into computer algorithms for pre-editing disease curves. Based on simulations it is shown that pre-editing time series of disability measured with the Expanded Disability Status Scale (EDSS) can lead to more robust and less biased estimates for important disease characteristics, such as baseline EDSS and time to reach certain EDSS levels or sustained progression.

1 Introduction

Multiple sclerosis (MS) is a complex autoimmune disease which is characterized by a great heterogeneity in disease course within as well as between individuals.¹ This results in uncertainty about diagnosis and prognosis. Natural history data providing information on the long-term course of MS patients are therefore especially valuable. They come from patient registries and studies other than clinical trials. Hence, natural history data are either collected according to a regular schedule or whenever a patient considered it necessary to consult his physician. Especially for the second group of studies it is quite likely that some of the data were collected while patients were affected by short-term symptoms of the disease, e.g. when suffering from a relapse. Results from former natural history studies and further references can be found, e.g. in articles by Weinshenker et al.^{2,3} and Confavreux et al.^{4,5}

Consequently, the variability of natural history data is not exclusively random but also contains a systematic component due to relapses. Not considering this systematic error would lead to biased estimates and predictions for the future course of the disease. There are, however, also other possible sources of systematic and random errors. In this paper interrater variability is also considered. It has been shown that especially for small EDSS values, different physicians arrive at different conclusions when diagnosing the same patient.⁶ Moreover, the complexity of the EDSS scale may imply the possibility of intrarater variability, i.e. that a single physician can come to different conclusions for patients with exactly the same symptoms.⁷ A typical source for a random error is the daily performance of a patient, i.e. a patient's EDSS value might change on a daily basis depending on, e.g. the temperature or also on the mood or the previous activity of a patient.⁸

To improve the reliability of estimates and predictions we suggest to apply neurological knowledge combined with ideas from the analysis of bio signals⁹ to filter out systematic effects that are caused by relapses. The other error sources then can be treated with statistical methods, like e.g. the segmented regression model proposed by.¹⁰ To evaluate the benefit of this filtering or pre-editing process the methods are first applied to simulated data which have the advantage that all parameters are known and then transferred to real data from the pooled database of the Sylvia Lawry Centre for Multiple Sclerosis Research (SLCMSR). Furthermore, we show by pre-editing data that come from placebo arms from clinical trials that

in definitions for sustained progression one has to draw the attention also on a stable baseline and not only on the assessment of ‘sustained’. The SLCMSR database contains anonymized data from natural history studies and placebo arms of clinical trials that have been donated by the pharmaceutical industry, by universities, hospitals and clinical research groups (donors: <http://www.slcmsr.org/en/partner.htm>).

2 Methods

2.1 Rules for detecting relapses

The rules for removing systematic errors due to EDSS values that seem too high, e.g. because of relapses, emerged from discussions among the authors who have neurological or mathematical background. Based on their clinical experience, neurologists identified outliers in EDSS courses drawn from the SLCMSR database. These outliers were then classified and simple rules derived that could be transformed into algorithms by the mathematicians. In addition to detecting these outliers the algorithms have to ensure that the basic structure of the data remains unaltered.

Basically, there are two rules. The first rule can be applied to both, clinical and natural history data, whereas the second rule is rather thought to be applied to natural history data only.

Assume the data set contains $i = 1, \dots, N$ patients and from each patient we have n_i EDSS measurements at times $t_{i1} < \dots < t_{in_i}$, where $t_{i1} = 0$ and t_{ij} is the elapsed time to the j^{th} observation in months (m). Formally, the rules modify the data set by a transformation:

$$\{EDSS(t_{ij})\} \longrightarrow \{EDSS'(t_{ij})\}.$$

In the following that transformation is described in detail. Since all operations are applied to each patient separately the subscript i is omitted in the following sections.

2.1.1 Rule 1 for high initial values

Patients are often suffering from short-term changes in their disability status, such as a relapse, when they first contact a physician. Hence this is also likely to be the case when they enter an MS registry or join a clinical trial. Consequently, the first observation, frequently used as a baseline value, may be artificially high. The first rule is, therefore, intended to filter the effect of a relapse at the initial visit. Since most patients with a relapsing remitting course of the disease recover from a relapse within three months, it seems appropriate to use the first observation recorded at least three months after entering the study as a revised baseline. In a few cases complete remission might take longer. In such a case the patient's EDSS levels would be expected to decrease over several months. If such a trend is observed the last value of this trend is used as the revised baseline. The initial value is set to the value of the revised baseline. Because of high correlation with the initial value also all values that are observed within the first three months are set to the revised baseline. This rule is only applied when the EDSS level at the entry visit is higher than at the following visits and possible adjustments are made before the checks for Rule 3, 4 and 5 are performed.

The mathematical formulation of this rule is given by:

For all j , such that $t_j < 3m$, the transformation is given by

$$EDSS'(t_j) = \min(EDSS(t_1), \dots, EDSS(t_l)), \quad (1)$$

where l is chosen such that $t_{l-1} \leq 3m$ and $t_l > 3m$.

If $EDSS(t_1) \geq \dots \geq EDSS(t_l)$ then an $l' \geq l$ is chosen such that $EDSS(t_1) \geq \dots \geq EDSS(t_{l'})$ and $EDSS(t_{l'}) < EDSS(t_{l'+1})$.

Finally, in equation (1) l is replaced by l' .

Figure 1 illustrates two possible applications of the initial value rule. No trend is observed in Figure 1a, however the first observation after three months has a smaller EDSS value than the initial one. Hence all values prior to three months are set to the value of the first observation after three months. In Figure 1b a trend is observed and the last value of the trend is used to modify the initial values.

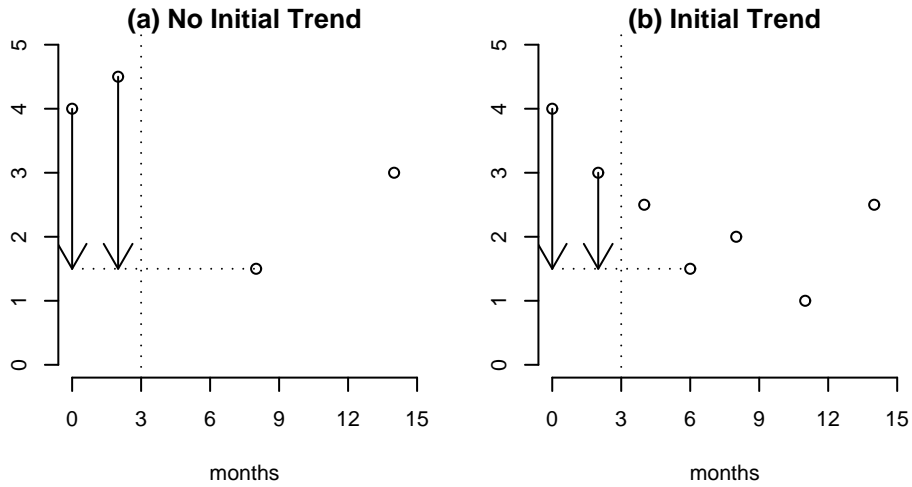


Figure 1: Rule for initial high values; The vertical line is drawn at three months.

2.1.2 Detecting outliers based on preceding observations

In this subsection the Rule 3 uses the information up to two years to identify an outlier. The basic idea is to recognize a trend in the EDSS course of a patient. Based on this trend an acceptance region for the following observation is defined. If the following value lies outside and its successor within this acceptance region the first value is regarded as unreliable and therefore discarded. Because of the high interrater variability for EDSS values smaller than or equal to 2.0, this rule is only applied for larger values. In contrast to Rules 1 and 2 that are applied immediately the transformations that result from this rule are only conducted after all checks including those for the later explained Rule 4 and 5 are conducted. This is to avoid a possible chain reaction that deletes all observed values from a patient and to ensure that decisions are based on as many observations as possible.

Before specifying this rule we need two definitions that explain what we mean by trend and acceptance region.

Trend A trend is specified by a linear regression line that is fitted for all points within a time window of 15 months. If there are less than two observations within this time window the immediately preceding values are used to include at least two points. If the time span for the points

used to evaluate the trend is less than or equal to three months also the preceding points are included until a period of at least three months is reached. A negative trend is not accepted. In this case the trend is given by a horizontal line through the last point that was used to compute the trend. Figure 2 illustrates this definition.

Acceptance region An acceptance region for an EDSS value observed at time t_j where $3 \leq j \leq n$ is defined by a trend line with time window $[t_j - 18m; t_j - 3m]$ and a slope parameter γ . Say the time point of the last observation that is used to compute the trend line is t_k . The upper boundary of the acceptance region is defined by a line. It is fixed by the value of the trend line at t_k and its slope is given by the sum of the slope of the trend line and γ . The area below the boundary is called the acceptance region. For a visualization of this definition see Figure 3.

With these definitions the rule is formulated as:

For all $j = 3, \dots, n - 1$ with $EDSS(t_j) \geq 2.5$ and $t_j > 6m$ a check is performed whether $EDSS(t_j)$ lies within its acceptance region with slope parameter $\gamma = 0.6 EDSS/year$. If it does not and $EDSS(t_{j+l})$ lies within the acceptance region, $EDSS(t_j)$ is deleted. Here l is the smallest integer such that $t_{j+l} - t_j > 3m$ or $EDSS(t_{j+l})$ is the last observation.

The slope parameter $0.6 EDSS/year$ corresponds to approximately double the value of the average increase in EDSS levels per year. This value was estimated based on the SLCMSR database. In Figure 3 the third observation would be deleted because it is outside the acceptance region and its successor lies inside.

2.1.3 Rule 4 for final values

Rule 3 cannot be applied to high final observations since no succeeding value is available. To compensate, the slope parameter for the acceptance region is increased to four times the average increase per year. Hence, the final value is deleted only in very rare and extreme cases.

The EDSS value $EDSS(t_n) \geq 2.5$ is deleted if it lies outside of its acceptance region with slope parameter $1.2 EDSS/year$.

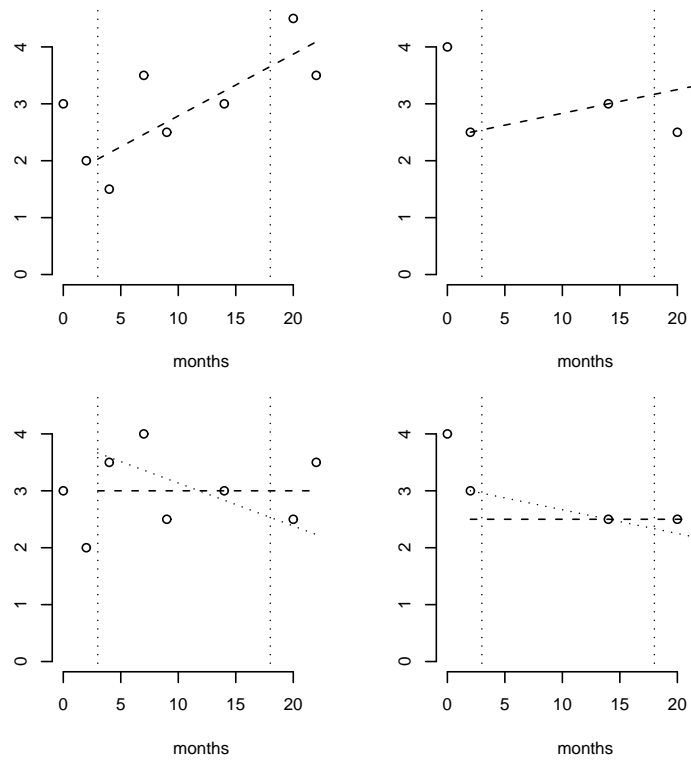


Figure 2: In all figures the trend line is specified by the dashed line. The dotted vertical lines indicate the 15 month time window. In the lower graphs the dotted lines are not accepted as trend lines because of their negative slopes and are replaced by the horizontal lines.

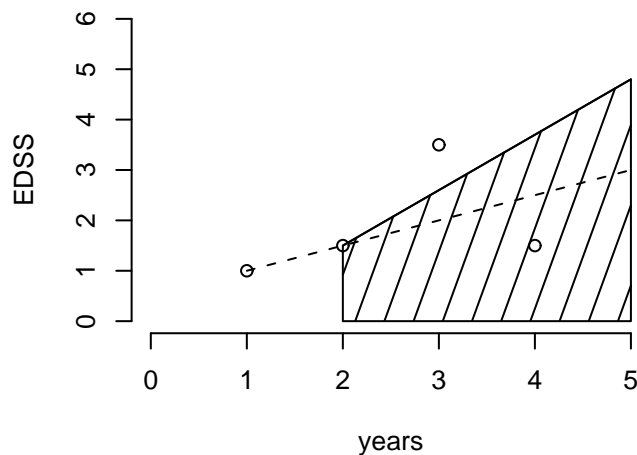


Figure 3: Illustration of an acceptance region. The dashed line indicates the trend line and the shaded polygon the acceptance region.

2.1.4 Rule 5 for exclusively small values

Basically, this rule does not preprocess the data but puts a very general restriction on the interpretation of time series of small EDSS values and on models to be fitted. It considers the high variability for EDSS values less than or equal to 2.0 as reversible fluctuations not contributing to disease progression. Furthermore, since these EDSS values are based on very small values on the different Kurtzke functional scales the interrater variability is rather high.^{6,11-13} Consequently, the rate or the starting point for the progression of MS should not be solely based on EDSS values that are smaller than 2.5.

If the maximal EDSS value is less than or equal to 2.0 then a progression has not begun yet.

Since these rules are supposed to be independent of the chosen model no more precise formulation can be given.

In the following section the impact of the rules on the parameters of a segmented regression model and on time to sustained progression is examined. Moreover a few graphs are shown that illustrate the influence of the rules on the data.

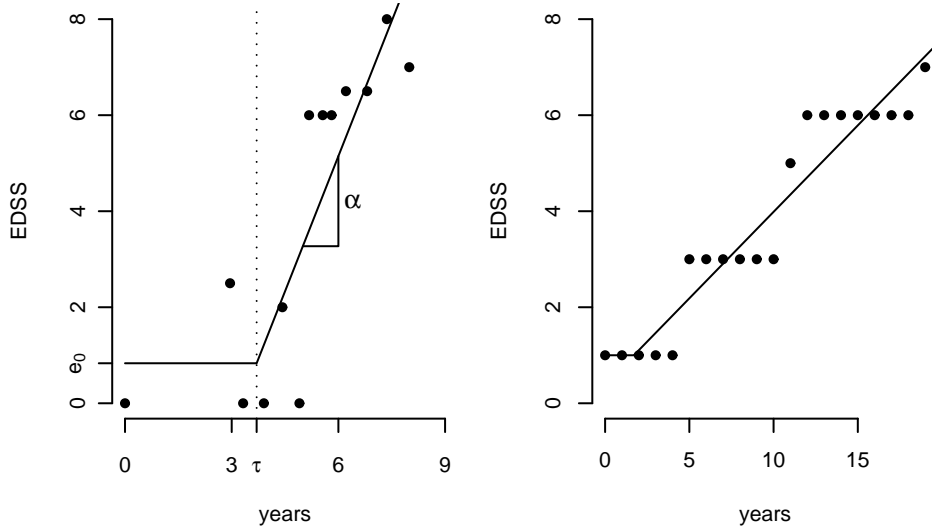


Figure 4: Fit of the segmented regression model for two real patients. In the graph on the left hand side, the baseline e_0 , the change point τ and the slope α are illustrated.

2.2 Simulation

In order to assess the effects of the introduced pre-editing rules on the parameter estimates of a segmented regression model data sets of simulated MS patients were generated from this model with and without added relapses.

The segmented regression model to be considered assumes that an initial stable phase of the disease is followed by a progressive phase and that the switch between them occurs at a specific time point (the “change point” τ). During the stable phase the observed EDSS values are assumed to fluctuate randomly around a fixed baseline value e_0 . After the change point the model assumes that the disease progresses linearly with slope α , i.e. the disease progression measured on the EDSS scale is constant at rate α . Either of the two phases might be missing. The three parameters are fitted for each patient separately. In Figure 4 the recorded EDSS values and the fitted segmented regression model are displayed for two patients.

The model can be described by¹⁰

$$EDSS(t_{ij}) = \begin{cases} e_{0i} + \epsilon_{ij} & \text{for } t_{ij} < \tau_i \\ e_{0i} + \alpha_i(t_{ij} - \tau_i) + \epsilon_{ij} & \text{for } t_{ij} \geq \tau_i \end{cases},$$

for $i = 1, \dots, N$ and $j = 1, \dots, n_i$, where e_{0i} refers to the baseline of the initial stable phase of the i^{th} patient, τ_i to the time until the change point, $\alpha_i > 0$ to the slope parameter of the regression line and ϵ_{ij} is some random noise. The parameters are fitted with a least squares approach and the restrictions arising from Rule 5. To ensure all three parameters can be estimated there cannot be a change point between the last but one and the last observation. For a more general discussion of segmented regression models and further references might be found in an article by Küchenhoff.¹⁴

For the simulation study 1000 patients, with one annual observation for ten years were generated. The underlying process was chosen to be the same for all patients: A segmented regression model with baseline $e_0 = 2$, change point $\tau = 4$ years and slope $\alpha = 0.5$ *EDSS/year*. Random noise was added to each observation by drawing from a discrete distribution with the following probabilities:

$$\begin{array}{c|ccccc} x & -1 & -0.5 & 0 & 0.5 & 1 \\ \hline P(X = x) & 0.1 & 0.25 & 0.3 & 0.25 & 0.1 \end{array}.$$

Additionally, it was assumed that a relapse in the data occurs independently at each time point with probability 0.2 and in such a case the EDSS level was increased by 3.

3 Results

In this section natural history data are shown before and after application of the rules to demonstrate their impact. Then their effect on the parameter estimates of the segmented regression model for simulated and real natural history data and on time to sustained progression in clinical data are examined.

The analyses are based on the SLCMSR data set `ver1_0_rc1` that was released in November 2003. This data set is split into an open and a closed part, where the closed part is only used for validation of hypotheses found in

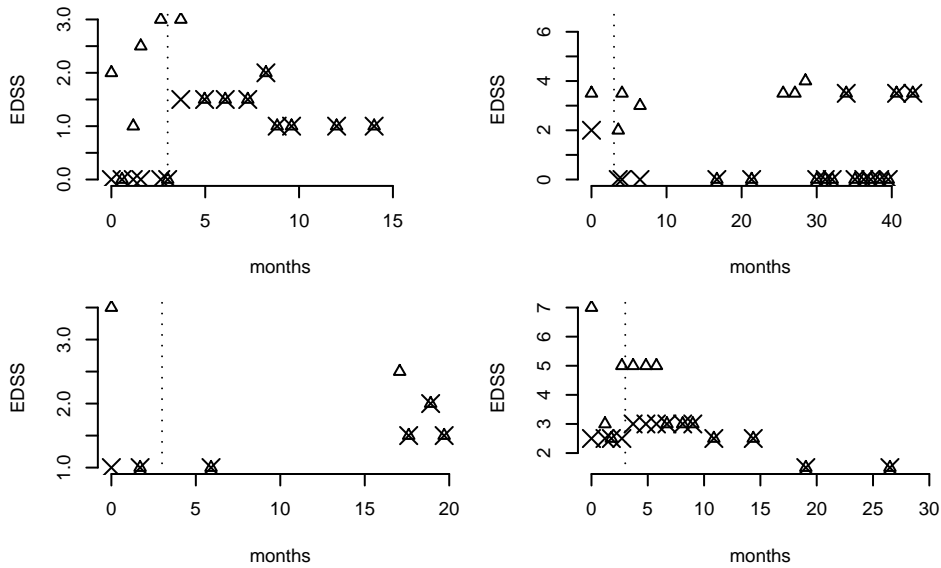


Figure 5: Disease course before (triangles) and after (crosses) application of the rules.

the open part; access to the closed part is limited to a data trustee. Hence only 40% of the data, that is 4781 natural history patients and 1354 placebo patients from clinical trials were available.

3.1 Pre-editing natural history data

To illustrate the impact of the rules we selected four patients from the natural history part of the SLCMSR database (see Figure 5). The patients have been chosen such that in all four graphs the first two rules have great impact: the initial values are reduced and the jagged courses are smoothed. The results for the initial value rules seem quite reasonable, except that in the upper right graph one might prefer to adjust the initial value to zero. On the other hand, a stricter Rule 1 would in many other cases lead to too substantial and undesirable changes in the data.

Rule 3 influences the EDSS course for the patients shown in the upper right and lower left graph. For the upper right graph, one might again be tempted to remove not only the three observations around month 28 but

also the observation around month 35 with EDSS level 3.5. This does not happen because the three observations are removed only after all observations are checked, i.e. the removed observations are still used to compute the trend line for the following observations.

When the rules were applied to natural history patients with at least five EDSS values in the SLCMSR data set (1120 patients with 10787 observations) the initial value rule was applied 244 times, Rule 2 199 times and Rule 3 and 4 lead to 566 deleted observations. The change in EDSS caused by Rule 1 and 2 was in 260 cases smaller than or equal to one, in 141 cases between one and three and in 42 cases greater than or equal to three.

3.2 Pre-editing and analysis of simulated natural history data

The effects of pre-editing on the parameter estimates of the segmented regression model are now examined. This is easiest, of course, when the true disease course is known. In order to achieve this data sets with EDSS time series of ten years were generated for 1000 patients using this model with and without added relapses. Subsequently, the segmented regression model was fitted to these simulated natural history data with and without first applying the rules.

In Table 1 the means and standard deviations for the parameter estimates are given. Because of the chosen simulated values all patients progress in both cases. Obviously, Rule 1 reduces the baseline, but also the estimate for the change point τ is underestimated. This effect is closely related to the lower baseline. In the case of an underestimated baseline the regression line for the progressive phase is lengthened until it touches the new baseline and hence leads to an earlier change point.

If the simulated time series include relapses, pre-editing leads to an estimate for the slope close to the simulated value, whereas without pre-editing the slope is estimated nearly twice as large as the simulated slope. Applying the rules to the data set without relapses increases the variability in the estimates but in the presence of relapses failure to apply the rules results in a very unreliable estimate for the slope.

The desirability of applying the rules becomes obvious when regarding the time until a certain EDSS level is reached. In Table 2 the mean times

Relapses	Rules	e_0		τ		α	
simu. values		2.00		4.00		0.50	
no	no	1.93	(0.30)	3.79	(1.46)	0.53	(0.16)
no	yes	1.68	(0.38)	3.43	(1.69)	0.53	(0.18)
yes	no	2.46	(0.59)	3.88	(2.58)	0.91	(0.91)
yes	yes	1.89	(0.55)	3.11	(2.21)	0.58	(0.36)

Table 1: Means (sd) for the parameter estimates for the segmented regression model. The model was fitted to each of the simulated data sets with and without relapses and with and without the application of pre-editing rules.

Relapses	Rules	EDSS 3		EDSS 5		EDSS 7	
simu. values		6.00		10.00		14.00	
no	no	6.08	(0.65)	10.10	(0.89)	14.11	(1.94)
no	yes	6.27	(0.67)	10.37	(1.13)	14.47	(2.30)
yes	no	4.01	(2.47)	9.13	(1.38)	13.20	(3.24)
yes	yes	5.29	(1.75)	9.87	(1.67)	14.20	(3.62)

Table 2: Estimates (sd) of time to landmark EDSS values for the simulated data sets.

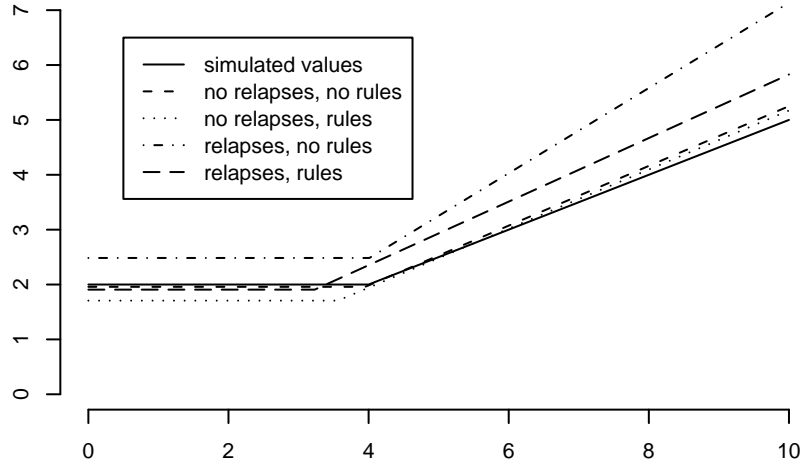


Figure 6: True and estimated EDSS curves based on the mean of the parameter estimates for the segmented regression model

until an EDSS level of three, five or seven is reached and the corresponding standard deviations are given. Applying the rules to the simulated data set without relapses leads to only slightly biased estimates but when relapses are present the estimation is improved considerably. Especially, for the time until an EDSS level of three is reached the naive estimates are highly biased and applying the rules reduces the bias significantly. However, by application of the rules in this case the variation increases.

The greatly improved ability of the model to predict certain landmarks of the disease can be seen in Figure 6. If no relapses are present the prediction with and without application of the rules differ only slightly. When relapses are present the application of the rules reduces the bias in possible predictions.

Hence the application of the rules will lead to a lower baseline and therefore to an earlier estimated starting point for the progression but when estimating the time of reaching certain landmarks in the disease, the application of the rules leads to less biased estimates. The magnitude of the biases depends mainly on the chosen parameters, however, the basic results can also

	e_0		τ in years		α in EDSS/year		no. prog.
w/o rules	2.25	(1.23)	4.20	(5.84)	0.463	(0.420)	39
w. rules	1.87	(1.29)	2.88	(4.69)	0.422	(0.386)	50
w. rules	2.00	(1.42)	3.54	(5.09)	0.496	(0.403)	39

Table 3: Parameter estimates (sd's) for the segmented regression model with and without application of the pre-editing rules. In the last row the estimates with application of the rules are given for the 39 patients that progress without application of the rules.

be obtained from simulations with different parameters, i.e. the qualitative nature of our conclusion does not depend on the chosen parameter of the simulation even though the quantitative aspects do.

3.3 Pre-editing and analysis of real natural history data

The segmented regression model is now applied to a subset of the natural history part of the SLCMSR data set with and without pre-editing of the EDSS curves. For a patient to enter the analysis, at least five visits and a duration, that is the time between onset of the disease and entry to the study, of less than 12 months were required. More than 350 of the eligible 422 patients belonged to studies with predominantly monotone disease courses, i.e. the EDSS levels for these patients increased or stayed constant between two visits; an example is displayed in the right graph in Figure 4. Since the rules have only little influence on such courses, a subset of 67 patients with variable course was obtained.

In Table 3 the mean parameters, their standard deviations and the number of patients who reach a progressing stage based on the segmented regression model are given. After application of the rules more patients reach

a progressing state. This is caused by Rules 1 and 2 that lead to a lower estimated baseline e_0 for many patients and hence a progression is more likely and in many cases an earlier starting point for the progressive phase is estimated. When the analysis is limited to the 39 patients that progress without applying the rules, the effects on the baseline and the change point are similar to those in the simulation. Both parameters decrease, however, the standard deviation for the baseline estimates becomes slightly larger.

3.4 Application of pre-editing for controlling errors in primary endpoints

The rules are primarily designed to be applied to natural history data, where not too many observations of a patient per year are available. In particular, the parameters for Rules 3 and 4, that depend on the definition of an acceptance region, would have to be changed if these rules are to be applied to clinical trial data that are recorded e.g. every three months. The initial value rule, however, is also applicable to clinical trial data.

In order to examine the influence of the first value on time to sustained progression, it was computed for placebo patients in the SLCMSR database with and without application of the initial value rule. For time to sustained progression a confirmation period of 180 days was chosen. That means a worsening was considered to be sustained if an EDSS value is observed that exceeds the initial value by at least 1 and in the following 180 days no smaller value is observed. If no observation is available in the following 180 days, the closest following value has to exceed the initial value by at least 1. Since a worsening in the upper half of the EDSS scale is considered to be more severe, for initial values that are greater than or equal to 6 the increase required is reduced to 0.5. Also, if there was no observation at least 180 days after exceeding the threshold, the worsening was not considered as sustained.

There are 1060 placebo patients with at least five visits in the SLCMSR database. Observation periods range from one to five years, where in most trials the patients were observed for two or three years. Measurements were taken at least every six months, in most cases every three months. For 371 patients a sustained progression is observed. After application of the initial value rule, however, a total of 464 patients reach the state of sustained progression. There are 367 patients that progress in both cases; for 274 of these

patients the times coincide. In 63 out of the 367 cases the application of the rules leads to shorter progression times and in 30 cases it was lengthened. Shorter progression times result from lower initial values; the average decrease for these cases is 2.8 months. Lengthened progression times arise from adjustment of values within the first three months, i.e. not only the initial value but also other values within the first three months are adjusted and a possible first exceeding of the threshold is delayed. The average increase for the 30 cases is 11.4 months.

Hence, if the baseline in the trials had been chosen to be the minimal EDSS value within the first three months, nearly an additional 10% of the placebo patients would have reached sustained progression. So far the focus discussion has focused on different definitions of ‘sustained’ with varying confirmation periods were discussed, but it seems necessary to turn one’s attention also to the best definition of the baseline value.

4 Discussion

The five rules presented here aim to reduce variability in EDSS data that is caused mainly by short-term changes of the disease course due to relapses. The approach is primarily intended for natural history data and is based on neurological knowledge and tries to formalize and simulate what many clinicians do when they interpret time series of disability measures. The careful pre-processing of disability curves using these rules is intended to improve estimation in modelling and thus our ability to predict the course of irreversible disability. The merits of this approach were investigated in various settings by comparing analyses of simulated or real disease courses with and without this pre-editing procedure.

The real strength of this approach became apparent when a segmented regression model assuming a change point between stable and progressive disease phases was used to determine the time until a certain disability level was reached. This analysis also revealed an interaction between the values for baseline disability (e_0) and change point (τ) estimated by this model which requires further investigation. Furthermore, the application of parts of the rules to placebo data from clinical trials illustrated the importance of baseline disability for definitions of time to sustained progression. Consequently, discussions of possible definitions of this end point should also concentrate

on establishing the most appropriate definition of a stable baseline.

Not applying such pre-editing rules might lead to overestimating the degree of irreversible disability. Their application might reduce this bias and improve the prediction of future EDSS levels. Although the rules were mainly investigated in combination with the segmented regression model, they are suitable for any scenario where relapses are likely to introduce unwanted noise in the data.

For different situations, e.g. equidistant time points between observations, it might be useful to adjust the parameters for the acceptance region for this situation. A further approach might be to define the acceptance region differently, e.g. one could think of limiting its upper boundary or perhaps change its slope for values above 6 as is done in some definitions for sustained progression.

In summary, the rules are recommended for use in situations where relapses are likely to be present and no information is available if an EDSS value was recorded during a relapse. The likely presence of relapses may be derived from knowledge of the study or simply by looking at individual EDSS courses. Since one hardly loses information when applying the rules in the absence of relapses but gains precision and achieves a bias reduction in their presence we propose to use this pre-editing tool.

Acknowledgements

The authors are grateful for careful reading and helpful comments from John Petkau. Ludwig Heigenhauser is funded by the SFB 386 of the German Science Foundation.

References

1. A. Compston, G. Ebers, H. Lassmann, I. McDonald, B. Matthews, and H. Wekerle. *McAlpine's Multiple Sclerosis*. Churchill Livingstone, 3 edition, 1998.
2. B.G. Weinshenker, B. Bass, G.P.A. Rice, J.H. Noseworthy, W. Carriere, J. Baskerville, and G.C. Ebers. The natural history of multiple sclerosis:

- a geographically based study. I. Clinical course and disability. *Brain*, 112:133–146, 1989.
3. B.G. Weinshenker, B. Bass, G.P.A. Rice, J.H. Noseworthy, W. Carriere, J. Baskerville, and G.C. Ebers. The natural history of multiple sclerosis: a geographically based study. II. Predictive value of the early clinical course. *Brain*, 112:1419–1428, 1989.
 4. C. Confavreux, S. Vukusic, and P. Adeleine. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. *Brain*, 126:770–782, 2003.
 5. C. Confavreux, S. Vukusic, T. Moreau, and P. Adeleine. Relapses and progression of disability in multiple sclerosis. *N Engl J Med*, 343:1430–1438, 2000.
 6. D.E. Goodkin, D. Cookfair, K. Wende, D.Bourdette, P. Pullicino, B. Scherokman, and R. Whitham. Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scales (EDSS). *Neurology*, 42:859–863, 1992.
 7. J. Hobart, J. Freeman, and A. Thompson. Kurtzke scales revisited: the application of psychometric methods to clinical intuition. *Brain*, 123:1027–1040, 2000.
 8. H. Albrecht, C. Wötzel, L.P. Erasmus, M. Kleinpeter, N. König, and W. Pöllmann. Day-to-day variability of maximum walking distance in MS patients can mislead to relevant changes in the Expanded Disability Status Scale (EDSS): average walking speed is a more constant parameter. *Mult Scler*, 7:105–109, 2001.
 9. M. Basseville and I.V. Nikiforov. *Detection of abrupt changes*. Prentice-Hall, 1993.
 10. B. Hellriegel, M. Daumer, and A. Neiß. Analysing the course of multiple sclerosis with segmented regression models. Technical report, Ludwig-Maximilians-University Munich, SFB Discussion Paper, 2003.
 11. M.P. Amato, L. Fratiglioni, C. Groppi, G. Siracusa, and L. Amaducci. Interrater reliability in assessing functional systems and disability on the Kurtzke scale in multiple sclerosis. *Arch Neurol*, 45:746–748, 1988.

12. D. M. Wingerchuk and B.G. Weinshenker. Multiple sclerosis: epidemiology, genetics, classification, natural history, and clinical outcome measures. *Neuroimaging Clin N Am*, 10:611–624, 2000.
13. J.H. Noseworthy, M.K. Vandervoort, C.J. Wong, and G.C. Ebers. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. The Canadian Cooperation MS Study Group. *Neurology*, 40:971–975, 1990.
14. H. Küchenhoff. An exact algorithm for estimating breakpoints in segmented linear models. *Computational Statistics*, 12:235–247, 1997.