



**Universiteit  
Leiden**  
The Netherlands

## **Measuring change in health-related quality of life: the impact of different analytical methods on the interpretation of treatment effects in glioma patients**

Coomans, M.B.; Taphoorn, M.J.B.; Aaronson, N.K.; Baumert, B.G.; Bent, M. van den; Bottomley, A.; ... ; EORTC Brain Tumor Grp

### **Citation**

Coomans, M. B., Taphoorn, M. J. B., Aaronson, N. K., Baumert, B. G., Bent, M. van den, Bottomley, A., ... Dirven, L. (2020). Measuring change in health-related quality of life: the impact of different analytical methods on the interpretation of treatment effects in glioma patients. *Neuro-Oncology Practice*, 7(6), 668-675. doi:10.1093/nop/npaa033

Version: Publisher's Version  
License: [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/)  
Downloaded from: <https://hdl.handle.net/1887/3184772>

**Note:** To cite this publication please use the final published version (if applicable).

## Measuring change in health-related quality of life: the impact of different analytical methods on the interpretation of treatment effects in glioma patients

Marijke B. Coomans, Martin J.B. Taphoorn, Neil K. Aaronson, Brigitta G. Baumert, Martin van den Bent, Andrew Bottomley, Alba A. Brandes, Olivier Chinot, Corneel Coens, Thierry Gorlia, Ulrich Herrlinger, Florence Keime-Guibert, Annika Malmström, Francesca Martinelli, Roger Stupp, Andrea Talacchi, Michael Weller, Wolfgang Wick, Jaap C. Reijneveld, and Linda Dirven; on behalf of the EORTC Quality of Life Group and the EORTC Brain Tumor Group

*Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands (M.B.C., M.J.B.T., L.D.); Department of Neurology, Haaglanden Medical Center, Den Haag, the Netherlands (M.J.B.T., L.D.); Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, the Netherlands (N.K.A.); Institute of Radiation-Oncology, Kantonsspital Graubünden, Chur, Switzerland (B.G.B.); Department of Radiation Oncology (MAASTRO clinic), and GROW (School for Oncology and Developmental Biology), Maastricht University Medical Center, Maastricht, the Netherlands (B.G.B.); The Brain Tumor Center, Erasmus MC Cancer Institute, Rotterdam, the Netherlands (M.V.D.B.); Quality of Life Department, European Organisation for Research and Treatment of Cancer, Brussels, Belgium (A.B., C.C., F.M.); Department of Medical Oncology, Azienda USL-IRCCS Institute of Neurological Sciences, Bologna, Italy (A.A.B.); Aix-Marseille Université, APHM, CNRS, INP, Inst Neurophysiopathol, CHU Timone, Service de Neuro-Oncologie, Marseille, France (O.C.); European Organization for Research and Treatment of Cancer Headquarters, Brussels, Belgium (T.G.); Division of Clinical Neurooncology, Department of Neurology, University of Bonn Medical Center, Bonn, Germany (U.H.); Groupe Hôpital Pitié-Salpêtrière, Paris, France (F.K.-G.); Department of Advanced Home Care and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden (A.M.); Northwestern University, Feinberg School of Medicine, Chicago, Illinois, US (R.S.); Department of Neurosciences, Azienda Ospedaliera San Giovanni Addolorata, Roma, Italia (A.T.); Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland (M.W.); Neurology Clinic and National Centre for Tumour Diseases, University Hospital Heidelberg, Heidelberg, Germany (W.W.); German Consortium of Translational Cancer Research (DKTK), Clinical Cooperation Unit Neurooncology, German Cancer Research Center, Heidelberg, Germany (W.W.); Department of Neurology and Brain Tumour Center Amsterdam, Amsterdam University Medical Center, Amsterdam, the Netherlands (J.C.R.); Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, the Netherlands (J.C.R.).*

**Corresponding Author:** Marijke Coomans, MSc, Leiden University Medical Center, Department of Neurology, PO Box 9600, 2300 RC Leiden, the Netherlands ([m.b.coomans@lumc.nl](mailto:m.b.coomans@lumc.nl)).

### Abstract

**Background.** Different analytical methods may lead to different conclusions about the impact of treatment on health-related quality of life (HRQoL). This study aimed to examine 3 different methods to evaluate change in HRQoL and to study whether these methods result in different conclusions.

**Methods.** HRQoL data from 15 randomized clinical trials were combined (CODAGLIO project). Change in HRQoL scores, measured with the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 and BN20 questionnaires, was analyzed in 3 ways: (1) at the group level, comparing mean changes in scale/item scores between treatment arms, (2) at the patient level per scale/item, calculating the percentage of patients that deteriorated, improved, or remained stable per scale/item, and (3) at the individual patient level, combining all scales/items.

**Results.** Baseline and first follow-up HRQoL data were available for 3727 patients. At the group scale/item level, only the item “hair loss” showed a significant and clinically relevant change (ie,  $\geq 10$  points) over time, whereas

change scores on the other scales/items were statistically significant only (all  $P < .001$ ; range in change score, 0.1-6.2). Although a large proportion of patients had stable HRQoL over time (range, 27%-84%) on the patient level per scale/item, many patients deteriorated (range, 6%-43%) or improved (range, 8%-32%) on a specific scale/item. At the individual patient level, the majority of patients (86%) showed both deterioration and improvement, whereas only 1% remained stable on all scales.

**Conclusions.** Different analytical methods of changes in HRQoL result in distinct conclusions of treatment effects, all of which may be relevant for informing clinical decision making.

### Keywords:

brain tumor | patient-reported outcome | quality of life | questionnaire

Functioning and well-being are particularly important for patients with an incurable disease such as glioma, for which both the duration and quality of survival count.<sup>1</sup> To quantify patients' functioning and well-being, health-related quality of life (HRQoL) questionnaires are often used. These questionnaires are typically multidimensional in nature, including single- and multi-item scales that assess functional health as well as symptom burden. Although these measures provide a lot of information about the functioning and well-being of patients, they also result in an analytical challenge because of the multiple outcomes that are generated.

In clinical trials, HRQoL scores are typically used to evaluate the impact of the treatments under investigation on the level of functioning and symptom burden on a group level. This means that changes in HRQoL scores over time are compared between treatment arms. In this case, HRQoL is assessed per scale/item, that is, at the "group scale/item level." Another way to look at a change in HRQoL scores is at the individual patient level. This can be achieved per scale/symptom (ie, "patient scale/item level") by calculating the percentage of patients whose HRQoL remained stable, improved, or deteriorated over time on a specific scale/item. However, for an individual patient, it may be of more interest to observe changes in the full range of scales/items simultaneously, rather than for only a single scale/item. Any given patient may improve on one scale/item, and deteriorate or remain stable on another scale/item. For example, a patient may remain stable in his or her level of physical functioning and pain during treatment, but may experience more fatigue.

To the best of our knowledge, there has only been one study in brain tumor patients that focused on investigating changes in HRQoL and whether conclusions on the impact of treatment on HRQoL differed when analyzed at the group or (individual) patient level.<sup>2</sup> In this small study, patients with brain metastases treated with stereotactic radiotherapy were found to have stable HRQoL scores over time when analyzed at the group level, but when analyzed at the individual patient level, many patients actually deteriorated or improved on specific HRQoL scales/items. More important, the majority of patients (64%) showed both improvement and deterioration on different HRQoL scales. Thus, different methods of analysis may result in different conclusions regarding treatment effects.

The aim of our study was to examine 3 different methods of evaluating change in HRQoL scores in a large group of glioma patients and to examine whether these methods result in different conclusions regarding the impact of treatment on HRQoL.

## Methods

### Study Sample

This study is part of the CODAGLIO (COmbining clinical DATatasets in GLIoma) project, in which a database was created including HRQoL data of individual glioma patients from 15 phase 2 and 3 randomized controlled trials (RCTs)(Supplementary Table 1). We included those RCTs in the database that assessed HRQoL with the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30 (QLQ-C30) and the complementary questionnaire for brain cancer patients QLQ-BN20.<sup>3</sup> HRQoL was assessed as a secondary end point in all RCTs. All patients gave their written informed consent to participate in the RCTs, and all principal investigators of these RCTs gave permission for use of the collected data.

### Health-Related Quality of Life Data

The EORTC QLQ-C30 consists of 5 functional scales (physical, role, emotional, cognitive, and social functioning), 3 symptom scales (fatigue, pain, and nausea and vomiting), a global health status/QoL scale, and 6 single items (dyspnea, appetite loss, insomnia, constipation, diarrhea, and financial difficulties). The QLQ-BN20 contains 20 items, comprising 4 symptom scales (future uncertainty, visual disorder, motor dysfunction, and communication deficit) and 7 single items (headaches, seizures, drowsiness, hair loss, itchy skin, weakness of legs, and bladder control). Raw scores for both questionnaires are linearly transformed to a scale from 0 to 100 according to the standard EORTC procedures.<sup>4</sup> For the functional scales and the global health status/QoL scale, a higher score indicates better HRQoL. For the symptom scales and items, higher scores indicate greater symptom burden. In all

RCTs, baseline questionnaires were administered before the start of the allocated treatment, but after surgery and irrespective of supportive treatment. To investigate changes in HRQoL, the first follow-up (FU) questionnaire scores of patients undergoing treatment were compared to their baseline scores. The timing of the first FU moment, reflecting the initial treatment effect, differed per RCT and ranged from 3 weeks to 16 weeks after baseline (median, 10.7 weeks). Clinically relevant change in HRQoL was defined as 10 or more points on a scale/item, reflecting the minimum clinically important difference.<sup>5</sup> For method 1, only those differences that were both statistically significant and clinically relevant were considered meaningful and therefore described. Methods 2 and 3 rely on determination of clinically relevant differences, and are therefore reported as such.

### Clinical and Sociodemographic Variables

The following available clinical and sociodemographic variables were collected: age, sex, tumor type (glioblastoma vs nonglioblastoma), prior surgery (resection vs biopsy), newly diagnosed vs recurrent tumor, World Health Organization (WHO) performance status (PS; 0 vs  $\geq 1$ ), and allocated treatment (radiotherapy, chemotherapy, angiogenesis inhibitors, tumor-treating fields, radiotherapy and chemotherapy combined, radiotherapy and angiogenesis inhibitors combined, radiotherapy combined with chemotherapy and angiogenesis inhibitors, and chemotherapy and TTF combined).

### Statistical Analysis

All patients with completed HRQoL baseline and FU questionnaires were included in the analysis. To evaluate whether there were differences between patients who completed both a baseline and FU questionnaire and patients who completed only baseline questionnaires, clinical characteristics were compared using the chi-square statistic for categorical data, and a *t* test for continuous variables.

#### Method 1: Change in Health-Related Quality of Life at the Group Scale/Item Level

Mean change scores for all HRQoL scales/items between baseline and the first FU assessment were calculated for all patients together, to identify significant and clinically relevant changes (ie,  $\geq 10$  points) at the group scale/item level.

#### Method 2: Change in Health-Related Quality of Life at the Individual Patient Level for Each Scale/Item

Changes in HRQoL scores were analyzed for each scale/item separately for each patient. To do so, differences in HRQoL scores between baseline and FU were computed for every scale/item. Thereafter, each patient was classified into 1 of 3 categories (improved, deteriorated, and stable) for every scale/item, based on the 10-or-more points

criterion for defining clinically relevant change. The percentage of patients in each category was computed.

#### Method 3: Change in Health-Related Quality of Life at the Individual Patient Level Including All Scales/Items

HRQoL scores were also analyzed at the individual patient level by considering all scales/items simultaneously. Based on the change in HRQoL scores on all scales/items, patients were categorized as (a) deteriorated, (b) improved, (c) stable, or (d) both improved and deteriorated. The “deteriorated” category included patients for whom the score of at least one scale/item declined and scores on the other scales/items remained stable. The “improved” category included patients for whom the score of at least one scale/item improved, and scores on the other scales/items remained stable. The “stable” category indicated stable, nonchanging scores on all scales/items. Last, the “declined/improved” category indicated that patients had deterioration in at least one scale/item and an improvement in at least one other scale/item. The distribution of patients in these 4 categories was visualized with heat maps. Patients were clustered not only based on their changes in the different HRQoL scales/items, but also on the clinical and sociodemographic variables age, WHO PS, sex, surgery, and tumor type, that represent the most distinct characteristics between patients in the included RCTs, to evaluate whether these factors were associated with a specific pattern of change in HRQoL scores. Analyses were performed using IBM SPSS, version 23.0.<sup>6</sup> The R packages Pheatmap<sup>7</sup> and ComplexHeatmaps<sup>8</sup> were used to create heat maps.

## Results

HRQoL data were available for 6084 patients in the CODAGLIO database, of whom 5217 patients (86%) completed an HRQoL baseline questionnaire, and 3727 patients completed both a baseline and first FU questionnaire (61%) (Table 1). Of the patients who completed a baseline and FU questionnaire, the majority were diagnosed with glioblastoma (68%), underwent resection (75%), and the mean (SD) age was 52 (13) years.

When compared with patients who did not complete an FU assessment, patients who completed a baseline and an FU questionnaire were younger (52 vs 53 years), less often diagnosed with glioblastoma (68% vs 71%), more often newly diagnosed patients (87% vs 83%), had a better PS (WHO = 0 in 43% vs 39%), underwent resection more often (75% vs 74%), and were more often allocated to a combination of treatments rather than monotherapy compared to the patients who completed a baseline questionnaire only, indicating minor imbalances (Table 1).

#### Method 1: Change in Health-Related Quality of Life at the Group Scale/Item Level

When calculating the mean change in HRQoL scores between baseline and FU for each scale/item separately (at the



**Table 1.** Clinical/Sociodemographic Characteristics of Patients With and Without a Health-Related Quality of Life Baseline and Follow-Up Questionnaire

	All patients (with and without HRQoL questionnaires) n = 6084	Patients with HRQoL baseline questionnaire only (A) n = 5217	Patients with both a baseline and FU questionnaires (B) n = 3727	Difference between(A) and (B), <i>P</i>
Male	3710 (61)	3211 (62)	2309 (62)	.337
Female	2351 (39)	2005(38)	1417 (38)	
Missing	23 (0)	1 (0)	1 (0)	
Age, mean, SD, y	53 (13)	53 (13)	52 (13)	< .001 <sup>a</sup>
Glioblastoma	4322 (71)	3716 (71)	2521 (68)	< .001 <sup>a</sup>
Nonglioblastoma	1762 (29)	1501 (29)	1206 (32)	
Newly diagnosed	4968 (82)	4330 (83)	3223 (87)	< .001 <sup>a</sup>
Recurrent	1116 (18)	887 (17)	504 (14)	
WHO PS 0	2257 (37)	2006 (39)	1595 (43)	< .001 <sup>a</sup>
WHO PS 1/2	3771 (62)	3191 (61)	2125 (57)	
WHO PS missing	56 (1)	20 (0)	7 (0)	
Resection	4379 (72)	3845 (74)	2807 (75)	< .001 <sup>a</sup>
Biopsy	1523 (25)	1221 (23)	798 (21)	
Missing	182 (3)	151 (3)	122 (3)	
TRT: radiotherapy alone	1349 (22)	1105 (21)	812 (22)	< .001 <sup>a</sup>
TRT: chemotherapy alone	1112 (18)	843 (16)	502 (14)	
TRT: angiogenesis inhibitor alone	126 (2)	106 (2)	80 (2)	
TRT: radiotherapy and chemotherapy	1633 (27)	1455 (28)	1153 (31)	
TRT: radiotherapy and chemotherapy and angiogenesis inhibitor	834 (14)	807 (16)	640 (17)	
TRT: chemotherapy and angiogenesis inhibitor	444 (7)	360 (7)	245 (7)	
TRT: chemotherapy and angiogenesis inhibitor	120 (2)	107 (2)	31 (1)	
TRT: tumor-treating fields alone	466 (8)	434 (8)	264 (7)	
TRT: chemotherapy and tumor-treating fields				

**Abbreviations:** FU, follow-up; HRQoL, health-related quality of life; WHO PS, World Health Organization performance status; TRT, allocated treatment.

Numbers in parentheses are percentages. *P* values are based on chi-square statistics.

<sup>a</sup>Indicates significant level *P* < .001.

group level), only the item “hair loss” showed a significant and clinically relevant change (ie,  $\geq 10$  points) over time, with a mean deterioration of 10.2 points. Change scores on the other scales/items showed a statistically significant (all *P* < .001) but not clinically relevant change, with changes in mean scores ranging between 0.1 and 6.2 points (Figure 1). These results suggest that the treatments under investigation did not have a clinically relevant impact on the level of functioning and wellbeing of the patients.

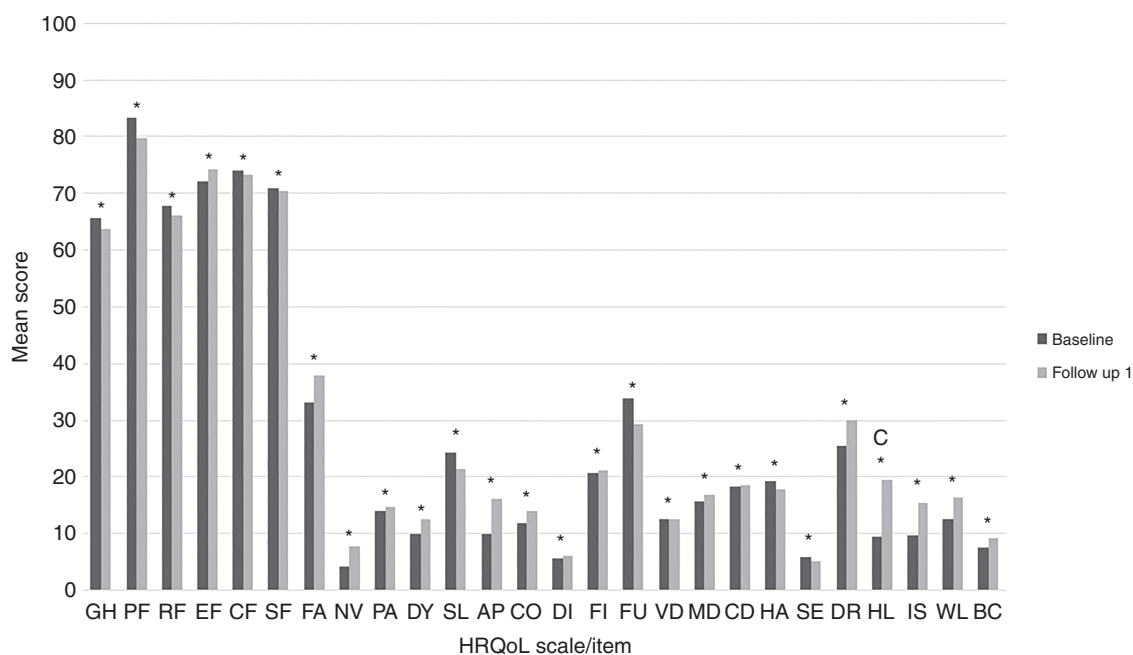
#### Method 2: Change in Health-Related Quality of Life at the Individual Patient Level for Each Scale/Item

Classification of patients into the 3 categories “stable,” “deteriorated,” and “improved” for each scale/item separately, showed that a large proportion of patients (range,

27%-84%) had “stable” scores on most items/scales. “Stable” was the largest category for all scales except fatigue, for which “deterioration” was the largest category. Nevertheless, Figure 2 also shows that a considerable percentage of patients had “deteriorated” (range, 6%-43%) or “improved” (range, 8%-32%) scores. Although the results at the group scale/item level showed that HRQoL was stable over time, the results of this analysis show that this does not hold true for a large proportion of patients.

#### Method 3: Change in Health-Related Quality of Life at the Individual Patient Level Including All Scales/Items

Analysis at the individual patient level considering all scales/items simultaneously showed that most patients



**Figure 1.** Health-Related Quality of Life (HRQoL) at the Group Scale/Item Level: Mean Scores for All HRQoL Scales/Items at Baseline and First Follow-Up Moment. AP, appetite loss; BC, bladder control; c, clinical relevant difference; CD, communication deficit; CF, cognitive functioning; CO, constipation; DI, diarrhea; DR, drowsiness; DY, dyspnea; EF, emotional functioning; FA, fatigue; FI, financial impact, FU, future uncertainty, GH, global health status; HA, headache; HL, hair loss; IS, itchy skin; MD, motor dysfunction; NV, nausea and vomiting; PA, pain; PF, physical functioning; RF, role functioning; SE, seizures; SF, social functioning, SL, insomnia; VD, visual disorder; WL, weakness of the legs. \*Statistically significant difference.

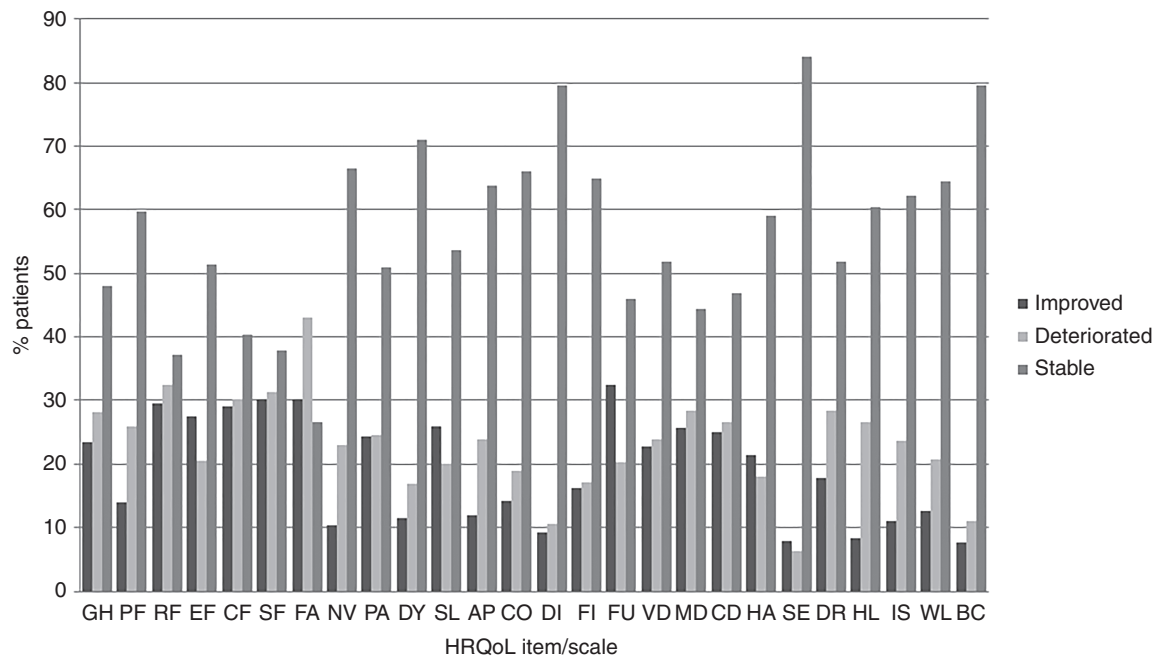
(86%) both deteriorated and improved, whereas only a minority of the patients only improved (6%), deteriorated (7%), or remained stable (1%) on all scales. In Figure 3, the change scores of all individual patients are visualized with a heat map. Similar to the analysis of individual patients at the individual scale/item level, this heat map demonstrates that within a patient the direction of HRQoL varies considerably. Indeed, the majority of patients both deteriorated on a scale/item, and improved or remained stable on other scales/items. Additional clustering on clinical characteristics (ie, WHO PS, sex, tumor type, surgery, newly diagnosed vs recurrent tumor, and age) did not identify subgroups of patients with a specific pattern of change in their HRQoL scores. Contrary to what one might expect, patients with more favorable characteristics, for example, younger age or better WHO PS, did not seem to be the patients who improved or remained stable with respect to their HRQoL scores.

## Discussion

The results of this study indicate that, depending on the method used to analyze and report HRQoL data, results may lead to different conclusions about treatment effects. When analyzing change in HRQoL scores of each scale/item at the group level, the results of this study suggest that initial treatment has hardly any clinically relevant impact on the functioning and well-being of glioma patients.

However, analyzing change in HRQoL scores at the individual patient level resulted in a different conclusion. First, although a large group of patients indeed remained stable on certain scales/items, an almost equal share of patients deteriorated or improved on those scales/items. This finding is masked when the data are analyzed only at the group level. A likely explanation may be that the scores for patients who deteriorated and improved averaged out, resulting in a stable score at the group level. Thus, averaging the scores for all patients together leads to the conclusion that there is no difference in HRQoL over time, whereas a significant percentage of the patients may, in fact, experience a decrease in their HRQoL. Furthermore, analyzing changes in HRQoL at the individual patient level including all scales/items simultaneously showed that the vast majority of patients (86%) both improved and deteriorated on different scales after treatment initiation, and only a small proportion of the patients remained completely stable over time (1%). This additional information about the joint impact of treatment on all outcomes may help patients and physicians to make the best treatment decision.

The patients included in this study appear to represent fairly well the population of glioma patients treated in clinical trials, which is of course already a selected group. Although patients who complete HRQoL questionnaires are a further selection of healthier patients,<sup>9</sup> those who completed both a baseline and FU questionnaire did not differ to a great extent from the patients who completed a baseline questionnaire only. Depending on the research question, different statistical techniques can be used. For

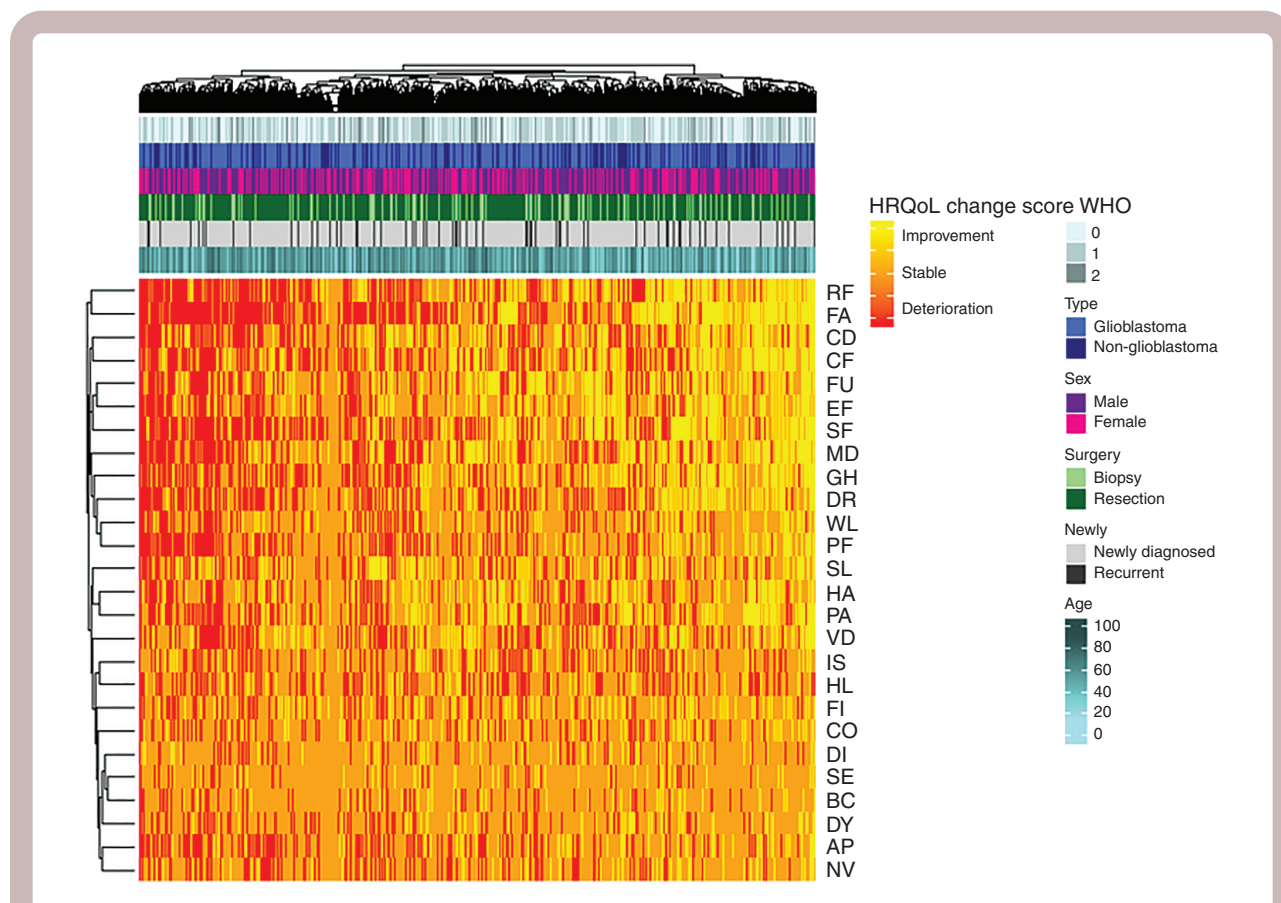


**Figure 2.** Health-Related Quality of Life (HRQoL) at the Individual Level Separately for Each HRQoL Scale/Item: Distribution of Patients Who Deteriorated, Improved, and Remained Stable Between Baseline and First Follow-Up Moment. AP, appetite loss; BC, bladder control; CD, communication deficit; CF, cognitive functioning; CO, constipation; DI, diarrhea; DR, drowsiness; DY, dyspnea; EF, emotional functioning; FA, fatigue; FI, financial impact; FU, future uncertainty; GH, global health status; HA, headache; HL, hair loss; IS, itchy skin; MD, motor dysfunction; NV, nausea and vomiting; PA, pain; PF, physical functioning; RF, role functioning; SE, seizures; SF, social functioning; SL, insomnia; VD, visual disorder; WL, weakness of the legs. Change in HRQoL scores were based on 10- or more point clinically relevant difference.

example, if one wants to analyze differences in HRQoL between 2 treatment arms over time, longitudinal models (eg, linear mixed models) should be used. In this study, we focused on a change in HRQoL between 2 time points, from baseline to the first follow-up assessment. This could also have been baseline and the assessment at the 12-month follow-up assessment. This again emphasizes that it is important to prespecify the research question with respect to HRQoL in the study protocol, and choose the appropriate statistical analyses accordingly. The choice of the HRQoL instruments may also affect our findings, particularly for the analysis at the individual patient level, because the QLQ-C30 and QLQ-BN20 together comprise 26 single- or multi-item scales. Indeed, the finding that only a tiny percentage of the patients (1%) remained completely stable over time in the simultaneous analysis of the individual patient level data including all scales/items can be explained by the large number of scales/items considered in the analysis as well as the definition of a clinically relevant change. It is unlikely that patients rate all 50 items in the EORTC QLQ-C30 and QLQ-BN20 exactly the same, even when their overall perceived HRQoL is unchanged. For single-item scales, this may directly result in a clinically relevant change of 10 or more points (eg, a change from “a little bit” to “quite a bit” will result in a change in item score of 33 points, thus exceeding the cutoff for a clinically relevant change). Currently, the EORTC Quality of Life group is working on a more sophisticated and appropriate way of defining clinically meaningful changes on the

scales/items of the QLQ-30 questionnaire, because recent studies highlight that the widely used 10-or-more points change<sup>5</sup> is too simplistic and does not detect a true clinical relevant change both at a group as well as at an individual patient level.<sup>10-12</sup> One method to reduce the impact of the abundance of scales/items would be to use a summary score for HRQoL, which is available for the EORTC QLQ-C30, integrating the majority of the functional and symptom scores.<sup>13</sup> A limitation, however, is that such a summary score including HRQoL issues relevant to brain tumor patients (ie, as measured with the QLQ-BN20) is currently not available. Ideally, analyses with this summary score should be performed at the group and at the individual patient level to extract maximum information with regard to change in HRQoL.

Our study shows that conclusions about the impact of treatment on HRQoL may vary depending on the analytical method used. We did not intend to extract information on the impact of specific treatments on HRQoL, nor for specific patient groups (eg, low-grade vs high-grade glioma), but focused on the impact of the analytical method chosen. When looking at individual clinical trials, however, it is important to evaluate the impact of treatment regimens. Results of past studies in glioma patients often have shown no difference in HRQoL comparing different treatment regimens, whereas, if analyzed at the individual patient level, important differences might have been observed.<sup>14-16</sup> As such, information given to patients could have been different. Possible clinically relevant



**Figure 3.** Health-Related Quality of Life (HRQoL) at the individual level taking into account all HRQoL scales/items: heat map reflecting change scores ( $\geq 10$ -point difference) for all HRQoL scales/items for all included patients. Patients and symptoms are ordered so similar scores are next to each other, using hierarchical clustering techniques. The horizontal axis represents all individual patients, and HRQoL scales/items are represented on the vertical axis. Red indicates deterioration in HRQoL, orange indicates stable scores, and yellow indicates improving scores. Annotations above the heat map indicate patients' clinical characteristics: World Health Organization performance status, tumor type, sex, surgery, newly diagnosed vs recurrent, and age. GH, global health status; PF, physical functioning; RF, role functioning; EF, emotional functioning; CF, cognitive functioning; SF, social functioning; FA, fatigue; NV, nausea and vomiting; PA, pain; DY, dyspnea; SL, insomnia; AP, appetite loss; CO, constipation; DI, diarrhea; FI, financial impact; FU, future uncertainty; VD, visual disorder; MD, motor dysfunction; CD, communication deficit; HA, headache; SE, seizures; DR, drowsiness; HL, hair loss; IS, itchy skin; WL, weakness of the legs; BC, bladder control. Change in HRQoL scores are based on 10- or more point clinically relevant difference.

information on the functioning and well-being of patients may be missed when analyzing HRQoL data at the group level only, because a large proportion of patients may experience a change in HRQoL that may go unnoticed at the group level if this proportion is similar, emphasizing the importance of analyzing HRQoL at the group and at the individual patient level. Moreover, although it is known that the first 2 methods described in our study result in different interpretations of outcomes, they are still not often both included in the study protocol and subsequently analyzed, reported, and interpreted as such. Therefore, we advise future trials to consider analyzing HRQoL data at the individual patient level in addition to the group level, and to prespecify these analyses in the study protocol. In addition, one could consider building a prognostic model, for example, to identify those patients who deteriorate during a specific treatment. Similarly, using a heat map to visualize and cluster HRQoL scores may also be incorporated into future research, because this adds to the interpretation on the impact of treatment on a patient's HRQoL.

Researchers or organizations may have their own procedures for the analysis and interpretation of HRQoL data. However, the diverse ways of analysis and interpretation of this data in the same clinical trial can result in conflicting and confusing conclusions, as also shown in this study. Different conclusions may of course be justified when different research questions are prioritized, but should not occur when answering the same research question. To address this issue, standardization of analytical methods with respect to HRQoL data are warranted. Currently, the Setting International Standards of Patient-Reported Outcomes and Quality of Life Endpoints Data in Cancer Clinical Trials initiative is ongoing, with the goal of establishing recommendations for the analysis of patient-reported outcomes in cancer clinical trials.<sup>17</sup> Ultimately, this guideline should improve the quality and consistency of statistical analysis in clinical trials, facilitating the interpretation of HRQoL findings.

In conclusion, when studying the impact of a treatment strategy on HRQoL in clinical trials, different analytical



methods may result in different conclusions about the impact of these treatments. Analyzing HRQoL at the individual patient level in addition to analysis of the scale/item at the group level may be valuable in providing insights, and should be considered in future research.

## Supplementary material

Supplementary material is available online at Neuro-Oncology Practice (<http://nop.oxfordjournals.org/>).

## Funding

This work was supported by a grant from the European Organisation for Research and Treatment of Cancer (EORTC Quality of Life Group (project number 1515).

## Acknowledgments

This study has been previously presented as an oral presentation at the European Association of Neuro-oncology (EANO) meeting September 20, 2019, in Lyon, France, and at SNO (Society for Neuro-Oncology annual meeting) November 22, 2019, in Phoenix, Arizona, USA. The abstract has been included in the abstract Supplementary in Neuro-Oncology: Coomans et al, OS7.2 Measuring change in health-related quality of life: the added value of analysis on the individual patient level in glioma patients in clinical decision making. *Neuro-Oncology*. 2019;21(suppl 3):iii14; <https://doi.org/10.1093/neuonc/noz126.045>.

**Conflict of interest statement.** A. Bottomley's and F. Martinelli's institution has received research funding from Boehringer-Ingelheim/Genentech/Merck, A. Malmström has received research funding from Cytovac Denmark, B. Baumert has received honoraria from Roche, and W. Wick has received research funding from Apogenix/Pfizer/Roche. All other authors declare no conflict of interest.

## References

1. Dirven L, Reijneveld JC, Taphoorn MJB. Health-related quality of life or quantity of life: a difficult trade-off in primary brain tumors?. *Semin Oncol*. 2014;41(4):541–552. doi:10.1053/j.seminoncol.2014.06.002.

2. van der Meer PB, Habets EJJ, Wiggeraad RG, et al. Individual changes in neurocognitive functioning and health-related quality of life in patients with brain oligometastases treated with stereotactic radiotherapy. *J Neurooncol*. 2018;139(2):359–368.
3. Taphoorn MJ, Claassens L, Aaronson NK, et al; EORTC Quality of Life Group, and Brain Cancer, NCIC and Radiotherapy Groups. An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-related quality of life and symptoms in brain cancer patients. *Eur J Cancer*. 2010;46(6):1033–1040.
4. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. *EORTC QLQ-C30 Scoring Manual*. Brussels, Belgium: EORTC Data Center; 2001.
5. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998;16(1):139–144.
6. IBM Corp. Released 2011. *IBM SPSS Statistics for Windows, Version 20*. Armonk, NY: IBM Corp; 2013.
7. Kolde R. Pheatmap: pretty heatmaps. *R package version [computer program]*. 2012:61.
8. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849.
9. Dirven L, Aaronson NK, Heimans JJ, Taphoorn MJB. Health-related quality of life in high-grade glioma patients. *Chin J Cancer*. 2014;33(1):40–45.
10. Musoro ZJ, Hamel JF, Ediebah DE, et al; EORTC Quality of Life Group. Establishing anchor-based minimally important differences (MID) with the EORTC quality-of-life measures: a meta-analysis protocol. *BMJ Open*. 2018;8(1):e019117.
11. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(2):171–184.
12. Giesinger JM, Kuijpers W, Young T, et al. Thresholds for clinical importance for four key domains of the EORTC QLQ-C30: physical functioning, emotional functioning, fatigue and pain. *Health Qual Life Outcomes*. 2016;14:87.
13. Giesinger JM, Kieffer JM, Fayers PM, et al; EORTC Quality of Life Group. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. *J Clin Epidemiol*. 2016;69:79–88.
14. Taphoorn MJB, Dirven L, Kanner AA, et al. Influence of treatment with tumor-treating fields on health-related quality of life of patients with newly diagnosed glioblastoma: a secondary analysis of a randomized clinical trial. *JAMA Oncol*. 2018;4(4):495–504.
15. Dirven L, van den Bent MJ, Bottomley A, et al; Dutch Neuro-Oncology Group (LWNO). The impact of bevacizumab on health-related quality of life in patients treated for recurrent glioblastoma: results of the randomised controlled phase 2 BELOB trial. *Eur J Cancer*. 2015;51(10):1321–1330.
16. Taphoorn MJ, Henriksson R, Bottomley A, et al. Health-related quality of life in a randomized Phase III study of bevacizumab, temozolomide, and radiotherapy in newly diagnosed glioblastoma. *J Clin Oncol*. 2015;33(19):2166–2175.
17. Bottomley A, Pe M, Sloan J, et al; Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) consortium. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. *Lancet Oncol*. 2016;17(11):e510–e514.