Calculating the net clinical benefit in neuro-oncology clinical trials using two methods: quality-adjusted survival effect sizes and joint modeling

Marijke B. Coomans, Linda Dirven, 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, Jeff A. Sloan, Roger Stupp, Andrea Talacchi, Michael Weller, Wolfgang Wick, Jaap C. Reijneveld, and Martin J. B. Taphoorn, 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., L.D., M.J.B.T.); Department of Neurology, Haaglanden Medical Center, Den Haag, the Netherlands (L.D., M.J.B.T.); 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.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 Univ, 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-Salpetrière, Paris, France (F.K.-G.); Division of Cell biology and Department of Biomedicine and Clinical Sciences, Linköping University, Linköping, Sweden (A.M.); Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA (J.A.S.); Northwestern University, Feinberg School of Medicine, Chicago, Illinois, USA (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; 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 B. Coomans, MSc, Department of Neurology, Leiden University Medical Center, PO BOX 9600, 2300 RC Leiden, the Netherlands (m.b.coomans@lumc.nl).

Abstract

Background. Two methods combining survival and health-related quality of life (HRQoL) data in glioma trials to calculate the "net clinical benefit" were evaluated: Quality-adjusted effect sizes (QASES) and joint modeling (JM).
Methods. The net clinical benefit in two trials was calculated as proof of concept for other trials. With the QASES method, effect sizes for differences in progression-free survival (PFS) or overall survival (OS) and HRQoL between the experimental arm and standard treatment arm were calculated, while the relative emphasis placed on survival/HRQoL varied. JM allows simultaneous modeling of HRQoL and OS/PFS.
Results. In the EORTC 26951 trial, combined radiochemotherapy significantly prolonged OS (difference 11.7 months), but also resulted in more patients experiencing clinically relevant worsening (≥10 points) in appetite loss and nausea/vomiting shortly after treatment. Using QASES, the survival benefit of additional procarbazine, lomustine, and vincristine (PCV) decreased from 42.3 months to 29.5 and 28.2 months when accounting for appetite loss and nausea/vomiting, respectively. JM analyses resulted in a loss of the beneficial effect of additional PCV between 13% and 24% when adjusting for different HRQoL parameters. The EORTC 22033 trial showed no

significant PFS difference between radiotherapy or temozolomide alone (46 vs 39 months), nor clinically relevant differences in HRQoL. JM analyses also showed no significant association between PFS and HRQoL scales/items, whereas QASES showed that temozolomide alone was more favorable when considering symptom burden (47–49 instead of 39 months).

Conclusions. Both methods resulted in different outcomes, but adjusting for the impact of treatment on HRQoL resulted in theoretically reduced survival benefits.

Key Points

- Two methods, joint modeling and quality-adjusted effect sizes are feasible to calculate the net clinical benefit.
- Both methods showed that adjusting for the impact of treatment on health-related quality of life resulted in theoretically reduced survival benefits.

Importance of the Study

This study used 2 distinct methods to calculate the "net clinical benefit" in 2 glioma clinical trials, combining survival and health-related quality of life (HRQoL) in one outcome. Both methods, joint modeling and quality-adjusted effect sizes, resulted in different outcomes, but are applicable to interpret HRQoL and survival data together. In clinical trials, these methods may facilitate interpretation on the net clinical benefit of a treatment strategy, while in clinical practice, they can be used to facilitate clinical decision making.

In clinical trials including glioma patients, data on health-related quality of life (HRQoL) are regularly collected in addition to traditional outcomes such as radiological tumor response, overall survival (OS), and progression-free survival (PFS). Incorporation of survival and quality-of-life outcomes into one expression seems important to properly evaluate treatment benefits. Indeed, a certain treatment strategy may prolong survival but may also be accompanied with considerable toxicity and subsequently a decreased level of functioning in patients. Hence, to determine the "net clinical benefit" of a new treatment strategy, both the quantity and quality of life should be considered.[1]

Until now, survival and HRQoL are often analyzed separately, and the lack of integration between the endpoints may lead to difficulties in interpretation of treatment benefits, particularly when results are conflicting. There are several statistical techniques available to combine survival and HRQoL outcomes into one expression. One method is joint modeling (JM), in which a longitudinal outcome (eg, HRQoL) and a survival outcome can be simultaneously analyzed. This technique has previously been applied to 2 trials in which oligodendroglioma patients were randomized to receive either radiotherapy (RT) alone or radiotherapy plus procarbazine, lomustine, and vincristine (PCV) chemotherapy.[2,3] Both trials showed that survival was still in favor of the combination treatment arm, even when adjusted for the negative impact of treatment on HRQoL.[2,3]

Another method to combine survival and toxicity information into a single expression of treatment benefit is with quality-adjusted effect sizes (QASES).[4] This method weighs differences in survival with differences in tumor response rate or toxicity, and one study previously applied this method to multiple trials including newly diagnosed glioblastoma patients[5]: The median PFS difference between two treatment arms (chemoradiation plus bevacizumab vs chemoradiation alone) decreased from 3.4 to 0.9 months in favor of the chemoradiation alone arm when adjusting for fatigue. Instead of toxicity, HRQoL data could also be used with the QASES method, thereby combining survival and HRQoL data.

The aim of this study was to apply both JM and QASES techniques to 2 different randomized controlled trials (RCTs) including glioma patients as a proof of concept for other trials, and evaluate differences between the methods.

Methods

Study Population

This study is part of the CODAGLIO (ie, *CO*mbining clinical trial *DA*tasets in *GLIO*ma) project, in which a database was created including HRQoL data of individual glioma patients from 15 previously published phase II/III RCTs. We calculated the net clinical benefit in 2 RCTs, EORTC 26951[6] (radiotherapy alone vs radiotherapy plus PCV chemotherapy in anaplastic oligodendroglioma and oligoastrocytoma) and 22033[7] (radiotherapy alone vs temozolomide alone in low-grade glioma), as a proof of concept for other trials, with 2 distinct methods.

HRQoL

HRQoL was assessed with the EORTC QLQ-C30 version 3.0 questionnaire,[8] and the EORTC brain cancer-specific QLQ-BN20.[9] Raw item scores were linearly transformed to scale scores ranging from 0 to 100 following the EORTC guidelines.[10] A clinically relevant change in HRQoL was defined as ≥10 points on a scale, reflecting the minimum clinically important difference.[11] Several HRQoL scales were preselected for analyses in this study (Supplementary File 1). Patients with HRQoL scores on at least one timepoint (baseline or follow-up) were included.

Joint modeling

Joint models allow simultaneous modeling of survival and a longitudinal endpoint,[12–14] for example, HRQoL measured over time. The joint model consists of 2 submodels: a longitudinal submodel (M1; HRQoL) and a time-to-event submodel (M2; survival), which are linked (joined) through a random effect.[13] For model 1 (M1), a flexible linear mixed effects model expanding the time effect into a B-splines basis matrix[3,15] was built using the longitudinal HRQoL outcome and included the treatment arm as covariate. For model 2 (M2), a Cox proportional hazards model was constructed including the time from randomization until the date of death (ie, event), or the date of last contact (ie, censored), and also the treatment arm as covariate. Finally, the 2 submodels M1 and M2 were linked through a random effect, resulting in the JM (ie, model 3 [M3]). In case of a significant association between survival and each HRQoL parameter (M1 and M2), outcomes of the JM including the effect of treatment on HRQoL (M3) were compared with outcomes of the Cox model (M2) including the effect of treatment on survival only. This way, an "adjusted" survival benefit of a treatment strategy was calculated.

First, we constructed JMs separately for each HRQoL scale. In a second step, in case HRQoL and survival were significantly associated in the univariable models, multivariate joint models were constructed, combining all selected HRQoL scales and survival into one model.[16] For the multivariate JM, patients with HRQoL scores on all selected scales available at least at one time point were included. Estimates in all models (M1–M3 and the multivariate JM) were adjusted for relevant prognostic sociodemographic and clinical factors including age, sex, WHO performance status (PS), and previous surgery (biopsy vs resection). Model assumptions were assessed graphically.[13] The formulas used to construct the models (M1-M3) are described in Supplementary Table 1.

QASES Method

A second method to combine HRQoL and survival into one outcome is with QASES, in which effect sizes for the treatment effect are first calculated separately for survival (ESS) and HRQoL (ESHRQoL). effect size survival (ESS) reflects the difference in mean survival time between the treatment arms (ie, survival time in the experimental arm minus survival time in the standard arm) divided by the standard deviation of the survival time in the standard arm (SDsurvival). ESHRQoL reflects the difference in the percentage of patients with a clinically relevant worsening (ie, ≥10 points at any timepoint during 1-year follow-up compared to the baseline HRQoL score) in the score on a preselected HRQoL scale between treatment arms, divided by the standard deviation of the sample proportion in that HRQoL scale of the standard arm (SDHRQoL).

Thereafter, QASES was calculated as the survival effect size minus the HRQoL effect size (ESS minus ESHRQoL). The resulting combined effect size can subsequently be weighted to vary the magnitude of the impact of survival and HRQoL. The weights for ESS (w1) and ESHRQoL (w2) can vary from 0 to 1, where w1 + w2 = 1. A weight of w1 = 1 indicates that survival is the only factor of importance, and HRQoL has no importance. Vice versa, a weight of w2 = 1 indicates that HRQoL is the only factor of importance. Similarly, w1 = 0.5 and w2 = 0.5 reflects the situation in which survival and HRQoL are considered equally important (see Supplementary Table 2 for all QASES equations). Ultimately, the combined effect size can be used to back-calculate the difference in survival between the treatment arms adjusted for the HRQoL effects.

All analyses were performed using IBM SPSS, version 23.0,[17] and R[18] using the JM package,[19] and the JMbayes package.[20]

Results

Results of EORTC trial 26951 showed a significant OS benefit with experimental treatment, but with significant and clinically relevant worse appetite loss.[13] For the EORTC 22033 trials, the PFS did not differ between treatment arms, nor did HRQoL.

Below we describe the results of the EORTC trial 26951 in full detail, to show how the 2 methods work, whereas the results of EORTC trial 22033 are explained only shortly, to highlight similarities and differences, but are described in full in Supplementary File 1.

Clinical Outcomes

The EORTC 26951 trial reported that OS was significantly longer in the RT/PCV arm (n = 182; 42.3 months) compared to the RT alone arm (n = 184; 30.6 months); HR = 0.75, 95% CI: 0.60–0.95.[6] Of the 366 included patients, 337 (92%) had baseline or follow-up HRQoL measures completed on at least one time point and were included in this study. There were no differences in clinical, sociodemographic, and molecular markers between patients with at least one HRQoL form and all patients (Table 1).

Clinical results of EORTC study 22033 showed no significant difference in PFS between radiotherapy only (46 months, n = 240) versus temozolomide chemotherapy only (39 months, n = 237) 1.16, 95% CI 0.9–1.5, P = .22).[7]

JM Outcomes

Results of the longitudinal sub models (M1) showed a worsening in functioning (ie, functional scales and

Table 1. Baseline Characteristics of the Included Patients in EORTC Trial 26951

Characteristics	All Patients (*n* = 366)	Patients With At Least One HRQoL Form (*n* = 337)	*P*Value
	n (%)	*n* (%)	
RT alone	184 (50)	171 (51)	.37
RT/PCV chemotherapy	182 (50)	166 (49)	
Male	210 (57)	192 (57)	.28
Female	156 (43)	145 (43)	
Age (mean, SD)	47.7 (11)	47.7 (11)	.11
WHO PS 0	134 (36)	124 (37)	.57
WHO PS 1	171 (47)	155 (46)	
WHO PS 2	58 (16)	55 (16)	
Missing	3 (1)	3 (1)	
Biopsy	52 (14)	47 (14)	.60
Partial resection	182 (50)	166 (49)	
Total resection	132 (36)	124 (37)	
MGMT: methylated	42 (11)	35 (10)	.46
MGMT: unmethylated	69 (19)	53 (16)	
MGMT: missing	255 (70)	249 (74)	
1p19q: codeleted	80 (22)	61 (18)	.44
1p19q: nondeleted	151 (41)	118 (35)	
1p19q: unknown	135 (37)	158 (47)	
IDH mutant	83 (23)	67 (20)	.09
IDH wildtype	98 (27)	81 (24)	
IDH missing	185 (50)	189 (56)	

HRQoL, health-related quality of life; IDH, isocitrate dehydrogenase mutations; MGMT, O^6-methylguanine-DNA methyltransferase promotor methylation; RT, radiotherapy; WHO PS, World Health Organization Performance Status.

the global health status) over time, and an increase in symptom burden (Table 2). Similar to the survival results reported for the entire EORTC 26951 population,[6] we found that combined RT/PCV was associated with prolonged OS in our subgroup with HRQoL data (M2; Table 1): HR = 0.83 (95% CI: 0.65–1.06), indicating a 17% lower risk of death for patients treated with RT/PCV compared to RT alone.

The JMs (M3) for the different HRQoL parameters showed a significant association (*P* < .0001) between OS and all preselected HRQoL scales except for nausea/vomiting, indicating that the risk for death was significantly increased with worsening global health, physical functioning, social functioning, and more fatigue, appetite loss, communication deficits, and seizures (Table 2). For example, a 10-point decrease in physical functioning corresponded to a 4% higher risk of death.

When comparing the survival models without HRQoL data (M2) with the JMs including HRQoL (M3 for separate HRQoL scales), we observed that the HRs for death in the JMs including HRQoL (HRs between 0.96 and 0.98 for the functional scales and between 1.01 and 1.03 for the symptom scales) were higher than in the survival model without HRQoL (HR = 0.83), accounting for a difference in HRs between 13% and 20% (Table 2). This means that, although the beneficial effect of combined treatment with

RT/PCV was still present, it was smaller when accounting for the negative impact of treatment on the patients' functioning and well-being.

Applying JM methods to EORTC trial 22033 showed no significant association between PFS and the HRQoL scales (Supplementary Table 5).

Multivariate Joint Model

Results of the JM including all preselected HRQoL scales for EORTC trial 26951 could be calculated for 270 patients that completed all HRQoL scales on at least one time point (Table 2). This multivariate model showed that only appetite loss and seizures were significantly associated with OS, and accounted for a theoretical loss of the treatment benefit of combined RT/PCV of 23% for both appetite loss and seizures. No multivariate JM was constructed for EORTC trial 22033 due to nonsignificant associations in the univariable JMs.

QASES Outcomes

The difference in OS time of 11.7 months between treatment arms in the EORTC 26951 trial resulted in an ESS of

Table 2. Comparison of the Cox Models (M2) and Joint Models (M3) in EORTC Trial 26951: Estimates Adjusted for the Clinical Variables Sex, Age, WHO Status, and Surgery

Model, Estimates	Estimate (SE)	HR (95% CI)	PValue	Difference in HR[c]
Univariate joint models (n = 332)				
M2: RT/PCV treatment (ref: RT alone)	–0.19 (0.12)	0.83 (0.65–1.06)	.12	
M3 joint models for selected HRQoL scales/items[a,b]				
RT/PCV treatment		0.76 (0.59–0.97)	.028*	
Global health status	–0.02 (0.01)	0.98 (0.97–0.99)	<.001**	0.15 (15%)
Physical functioning	–0.04 (0.01)	0.96 (0.94–0.98)	<.001**	0.13 (13%)
Social functioning	–0.02 (0.01)	0.98 (0.98–0.99)	<.001**	0.15 (15%)
Fatigue	0.02 (0.01)	1.02 (1.01–1.03)	<.001**	0.19 (18%)
Appetite loss	0.02 (0.01)	1.02 (1.01–1.03)	.0001**	0.19 (19%)
Communication deficit	0.01 (0.01)	1.01 (1.00–1.02)	.018*	0.18 (18%)
Seizures	0.03 (0.1)	1.03 (1.02–1.05)	<.001**	0.20 (20%)
Nausea/vomiting	0.01 (0.02)	1.01 (1.00–1.02)	.76	NA
Multivariate joint model (n = 270)				
M2: RT/PCV treatment (ref: RT alone)		0.80 (0.62–1.03)	.08	
M3: Multivariate joint model				
RT/PCV treatment	–0.276 (0.00)	0.76	.054	NA
Global health status	0.002 (.00)	1.00	.86	NA
Social functioning	–0.010 (0.00)	0.99	.23	NA
Fatigue	–0.005 (0.00)	0.99	.58	NA
Appetite loss	0.031 (0.00)	1.03	<.0001**	0.23 (23%)
Communication deficit	–0.006 (0.00)	0.99	.11	NA
Seizures	0.026 (0.00)	1.03	<.0001**	0.23 (23%)
Nausea/vomiting	0.006 (0.00)	0.94	.054	NA

Positive estimates indicate that scores on scales/items increase over time, meaning better scores for the functional scales and global health status and worse scores on the symptom scales, whereas negative estimates indicate that scores on scale/items decreased over time, meaning worse scores for the functional scales and global health status and better scores on the symptom scales.
HR, hazard ratio for the event (death); HRQoL, health-related quality of life; M2, cox model including the treatment effect; M3, joint model; SE, standard error; NA, not applicable in case of nonsignificance; RT, radiotherapy.
[a]Reference group: radiotherapy alone; [b]For the global health status, physical and social functioning: higher scores indicate better functioning. For the symptoms fatigue, appetite loss, communication deficit, seizures and nausea/vomiting, a higher score indicates more symptoms; [c] Difference in the HR between the cox model (M2) and the joint models (M3).
*Statistically significant at the <.05 level; **statistically significant at the <.01 level.

0.27 (Table 3), which can be considered a small effect size.[21] For HRQoL, the percentage of patients experiencing a clinically relevant deterioration on a HRQoL scale ranged from 15% to 41% in the RT alone arm, and from 10% to 42% in the RT/PCV arm (Table 3). In line with previous results,[22] a higher percentage of patients in the RT/PCV experienced symptoms during treatment: 47% of the patients in the RT/PCV arm experienced appetite loss versus 33% in the RT alone arm, and 39% versus 23% of patients experienced nausea and vomiting, respectively. Based on these percentages, the calculated ESHRQoL was 0.31 for appetite loss and 0.37 for nausea and vomiting. Next, the survival effect size was adjusted for these HRQoL effect sizes. With equal weights for survival and HRQoL (w1 = w2 = 0.5), the combined effect size was –0.02 for appetite loss and –0.05 for nausea/vomiting, indicating that including both HRQoL

and survival resulted in an unfavorable outcome for the RT/PCV arm. For appetite loss, this was equivalent to a median quality-adjusted OS for the RT/PCV arm of 29.5 months (instead of 42.3 months) compared to 30.6 months in the RT alone arm. Similarly, for nausea and vomiting, the median quality-adjusted OS was 28.2 months. For all other HRQoL scales, applying the same weights (w1 = w2 = 0.5), the median OS difference remained in favor of the RT/PCV arm. Figure 1 demonstrates the range of quality-adjusted survival differences for each preselected HRQoL scale with varying weights for survival and HRQoL.

In short, results of EORTC trial 22033 showed that the 7 months PFS benefit in the RT alone arm was nonsignificant. Nevertheless, a higher proportion of patients in the RT alone arm experienced worse social functioning, and more communication deficits, visual

Table 3. Summary of EORTC Trial 26951 Endpoints and Quality-Adjusted Effect Sizes

Endpoint	% of Patients Who Experienced a Clinically Relevant Deterioration During Treatment and Follow-up[a]		Difference in OS and HRQoL Between Arms	Effect Size HRQoL (ESHRQOL)	Total Effect Size (w1 + w2 = 1)	Back Calculated Median Difference	Adjusted Survival Time for Treatment With RT Alone[b]
	RT/PCV	RT alone					
Median OS (months)	42.3	30.6	11.7				
HRQoL scales/items							
Physical functioning	14.3	16.3	−2	−0.05	0.16	7.05	37.65
Social functioning	29.1	31.5	−2.4	−0.05	0.16	6.99	37.59
Global health status	29.1	25.5	3.6	0.08	0.09	4.03	34.63
Fatigue	41.2	40.8	0.4	0.01	0.13	5.67	36.27
Appetite loss	47.3	32.6[a]	14.7	0.31	−0.02	−1.07	29.53
Communication deficit	34.1	33.2	0.9	0.02	0.12	5.43	36.03
Seizures	9.9	15.2	−5.3	−0.15	0.21	9.11	39.71
Nausea/vomiting	38.5	22.8[a]	15.7	0.37	−0.05	−2.41	28.19

For appetite loss and nausea and vomiting, the back calculated median difference is negative, indicating that the mean survival difference between the treatment arms is in favor of the standard treatment arm (RT alone), instead of the experimental treatment arm (RT+PCV).

HRQoL, health-related quality of life; OS, overall survival; RT, radiotherapy.

[a]Percentage of patients with a clinically relevant deterioration in the experimental treatment arm compared to the standard treatment arm.

[b]Median OS in the RT alone arm plus the back calculated median difference (eg, 30.6–1.07 for appetite loss is 29.53 months).

disorders, motor dysfunction, and drowsiness, while in the TMZ arm, a higher proportion of patients experienced a clinically relevant decrease in global health status and role functioning. Including the impact of treatment on HRQoL in the analysis resulted in a larger adjusted PFS difference between treatment arms when adjusting for the global health status and role functioning (37–38 months instead of 39 months for TMZ alone compared to 46 months for RT alone), and an opposite effect when adjusting for communication deficit, visual disorder, motor dysfunction, and drowsiness (47–49 months instead of 39 months for TMZ alone compared to 46 months RT alone). However, these differences would most likely still not have resulted in a statistically significant difference between the treatment arms (Supplementary File 1).

Discussion

This study aimed to apply 2 methods, JM and QASES, to combine survival and HRQoL data into one outcome to facilitate interpretation of the net clinical benefit of a treatment strategy. In this study, we focused on 2 EORTC trials, as a proof of concept for other clinical trials. The EORTC 26951 showed that patients had a significant OS benefit (11.7 months) when treated with PCV chemotherapy in addition to RT compared to RT alone, but also a deterioration in appetite loss. In previous reports, these outcomes were analyzed and interpreted separately[22]: although combined treatment resulted in prolonged OS, there was a transient negative impact on HRQoL. Combining survival and longitudinally collected HRQoL data into one outcome confirmed that combined PCV/RT resulted in prolonged OS compared to RT alone, but this benefit was less pronounced when corrected for the negative impact of the treatment on HRQoL parameters. Results of EORTC trial 22033, in which no significant PFS benefit nor a clinically relevant impact of treatment on HRQoL on group level was found, also showed that when combining PFS and HRQoL into one outcome, the PFS difference between the treatment arms changed. Even when no significant and clinically relevant differences are found between treatment arms for survival and HRQoL when analyzed independently, joint analyses could lead to a different interpretation and subsequently different treatment decisions. Although both methods showed that adjusting for the negative impact of treatment on several HRQoL parameters theoretically reduced the survival benefit of experimental treatment versus standard treatment, the results were different. In EORTC trial 26951, one major difference is that the JMs, irrespective of the scale assessed, resulted in a survival benefit in favor of the RT/PCV chemotherapy treatment arm, whereas with the QASES method the survival benefit was in favor of the RT alone arm when considering certain scales (ie, appetite loss). This result is inherent to the statistical techniques, whereas JM uses longitudinally measured data, QASES is a cross-sectional method including the proportion of patients with deteriorating HRQoL scores as observed during treatment. This was also evident with the nausea/vomiting scale of EORTC trial 26951 which was not found to be associated with survival in the JM, while the QASES results resulted in an unfavorable outcome for the RT/PCV arm when including nausea and vomiting, similar to the original trial results.[13]

Another difference is that JM relies on statistical techniques that require statistical significance, whereas QASES
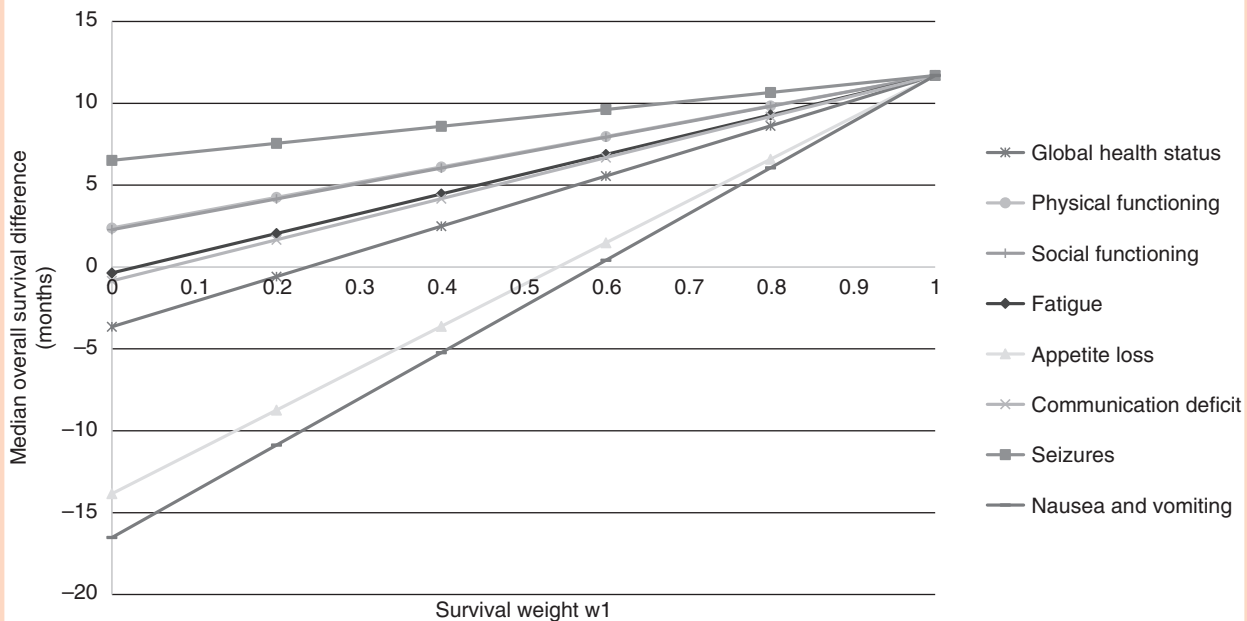
Figure 1. Quality-adjusted survival difference with varying weights of EORTC trial 26951, weights for survival (w1) and the health-related quality of life (HRQoL) scale/item (w2). W1 = 1 (right side of the figure), represents the hypothetical situation in which the patient considers survival as the only criteria of importance (survival: w1 = 1, HRQoL: w2 = 0). The left side of the figure represents the hypothetical situation in which the patient considers HRQoL as the only criteria of importance (w1 = 0, w2 = 1). The vertical axis indicates the median overall survival difference, with positive values representing a survival benefit in favor of the experimental treatment arm (RT/PCV) and the negative values indicating a survival benefit in favor of the standard treatment arm (RT alone).

does not rely on statistical significance but rather calculated effect sizes based on proportion. In case of the EORTC 22033 trial, JMs resulted in nonsignificant associations between survival and HRQoL and thus provided no information, whereas the net clinical benefit using QASES could be calculated.

Which method is preferred depends on the specific research question, and advantages and disadvantages of the models should be taken into account. For JM, the longitudinal and survival models where JM is based on have important characteristics such as the possibility to adjust for important prognostic clinical/sociodemographic factors, and the possibility to fit the model to the data over time. Importantly, treatment may result in toxicity on both the short and long term, and could be transient and/or persistent. The longitudinal nature of JM includes all HRQoL follow-up data available, and the interpretation of the net clinical benefit therefore depends depending on the duration of the follow-up data. Also, JM can handle missing data, a known challenge with HRQoL data.[23] Another advantage of JMs is that the results of these models can be used for monitoring in clinical practice, as dynamic predictions of an individual patients' survival probability can be produced, updated with ongoing measurements of an individual patient's HRQoL.[3] One disadvantage of JM may be the complex statistical analysis, and the dependency of the analysis on several strict model assumptions (eg, missingness at random of longitudinal data).

An advantage of the QASES method is that it is an intuitive and relatively easily applicable method, for which

only the percentage of patients with a deterioration in a specific HRQoL scale is needed. Also, the method allows to incorporate the patient's view on the relative importance ascribed to both quantity and quality of life, by adjusting the weights assigned to survival and HRQoL, thereby making this method meaningful for shared decision making. For example, in EORTC trial 26951, if a patient particularly worries about nausea and vomiting caused by a specific treatment strategy (eg, combined RT and PCV), and less about the length of survival, RT alone may be preferred. In contrast, for another patient, these symptoms may be of less importance while the length of survival is, resulting in a preference for combined RT/PCV.

Our study showed that it is feasible to use both JM and QASES in future clinical trials to interpret HRQoL and OS data together. This facilitates interpretation on the net clinical benefit of a treatment strategy. Combined interpretation of outcomes may be particularly relevant for personalized treatment and shared decision making, as it allows more effective communication between patients and physicians about the impact of a treatment strategy. In our view, the most important difference between these techniques is the inclusion of the number of measurements. QASES uses a change in HRQoL scores between 2 different time points, for example, between baseline and end of treatment, depending on the specific research question. JM on the other hand, makes use of all measurements over time (often more than 2), reflecting the complete longitudinal course of HRQoL. Future studies should investigate whether physicians find the use of these methods

helpful in determining the net clinical treatment benefit in clinical practice. The results of these models should then be presented in such a way that they are easily interpretable for patients and physicians. One example to present results of JMs to patients is by using dynamic predictions, which are useful to visualize the combined impact of treatment on the patients' HRQoL and prognosis. Dynamic predictions show the patients' survival probabilities in a graph which is updated when new HRQoL data is available.[3,24]

In conclusion, JM and the QASES method can be used to combine survival and HRQoL data into one outcome in glioma trials, but outcomes may differ. In clinical trials, these methods may facilitate interpretation on the net clinical benefit of the treatment under investigation, while in clinical practice they can be used to facilitate clinical decision making, for example via a computerized decision-making application.

Supplementary Material

Supplementary material is available at *Neuro-Oncology Advances* online.

Keywords

glioma | joint model | net clinical benefit | quality of life | survival

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. Epub 2014 Jun 11. PMID: 25173146.

2. Wang M, Cairncross G, Shaw E, et al.; Radiation Therapy Oncology Group (RTOG); North Central Cancer Treatment Group (NCCTG); Southwest Oncology Group (SWOG); National Cancer Institute of Canada Clinical Trials Group (NCIC CTG); Eastern Cooperative Oncology Group (ECOG). Cognition and quality of life after chemotherapy plus radiotherapy (RT) vs. RT for pure and mixed anaplastic oligodendrogliomas: radiation therapy oncology group trial 9402. *Int J Radiat Oncol Biol Phys.* 2010;77(3):662–669.

3. Ediebah DE, Galindo-Garre F, Uitdehaag BM, et al. Joint modeling of longitudinal health-related quality of life data and survival. *Qual Life Res.* 2015;24(4):795–804.

4. Sloan JA, Major B, Novotny PJ, et al. Combining survival and toxicity effect sizes from clinical trials into an interpretable, quality-adjusted survival effect size estimate of treatment efficacy. *J Clin Oncol.* 2014;32(15_suppl):6630–6630. doi:10.1200/jco.2014.32.15_suppl.6630.

5. Sloan J, Major B, Buckner JC. Qlif-08. Determining the net clinical benefit in Neuro-oncology clinical trials by combining survival and toxicity data. *Neuro Oncol.* 2017;19(Suppl 6):vi202.

6. van den Bent MJ, Brandes AA, Taphoorn MJ, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. *J Clin Oncol.* 2013;31(3):344–350.

7. Baumert BG, Hegi ME, van den Bent MJ, et al. Temozolomide chemotherapy versus radiotherapy in high-risk low-grade glioma (EORTC 22033-26033): a randomised, open-label, phase 3 intergroup study. *Lancet Oncol.* 2016;17(11):1521–1532.

8. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* 1993;85(5):365–376.

9. 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.

10. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. On behalf of the EORTC Quality of Life Group. *The EORTC QLQ-C30 Scoring Manual* 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.

11. 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.

12. Little RJ, Rubin DB. *Statistical Analysis with Missing Data.* Vol 793. Hoboken, New Jersey: John Wiley & Sons; 2019.

13. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. Boca Raton, Florida: Chapman and Hall/CRC; 2012.

14. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *J Clin Oncol.* 2010;28(16):2796–2801.

15. De Boor C, De Boor C, Mathématicien E-U, De Boor C, De Boor C. *A Practical Guide to Splines.* Vol 27. New York, NY: Springer-Verlag; 1978.

16. Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat Med.* 2011;30(12):1366–1380.

17. IBM Corp. Released. 2015. *IBM SPSS Statistics for Windows, Version 23.0.* Armonk, NY: IBM Corp; 2016.

18. Team RC. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing, 2013. http://www.R-project.org/.

19. Rizopoulos D. JM: an R package for the joint modelling of longitudinal and time-to-event data. *J Stat Softw.* 2010;35(9):1–33.

20. Rizopoulos D. The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. *arXiv:1404.7625.* 2014.

21. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*: Academic press; 2013.

22. Taphoorn MJ, van den Bent MJ, Mauer ME, et al. Health-related quality of life in patients treated for anaplastic oligodendroglioma with adjuvant chemotherapy: results of a European Organisation for Research and Treatment of Cancer randomized clinical trial. *J Clin Oncol.* 2007;25(36):5723–5730.

23. Dirven L, Aaronson NK, Heimans JJ, Taphoorn MJ. Health-related quality of life in high-grade glioma patients. *Chin J cancer.* 2014;33(1):40–45.

24. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics.* 2011;67(3):819–829.