A European Organisation for Research and Treatment of Cancer phase III trial of adjuvant whole-brain radiotherapy versus observation in patients with one to three brain metastases from solid tumors after surgical resection or radiosurgery: quality-of-life results.

This is the author's manuscript

Original Citation:

Availability:
This version is available http://hdl.handle.net/2318/135900 since

Published version:
DOI:10.1200/JCO.2011.41.0639

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)
This is an author version of the contribution published on:


The definitive version is available at:

http://jco.ascopubs.org/content/31/1/65.long
A European Organisation for Research and Treatment of Cancer Phase III Trial of Adjuvant Whole-Brain Radiotherapy Versus Observation in Patients With One to Three Brain Metastases From Solid Tumors After Surgical Resection or Radiosurgery: Quality-of-Life Results

Riccardo Soffietti, Martin Kocher, Ufuk M. Abacioglu, Salvador Villa, François Fauchon, Brigitta G. Baumert, Laura Fariselli, Tzahala Tzuk-Shina, Rolf-Dieter Kortmann, Christian Carrie, Mohamed Ben Hassel, Mauri Kouri, Egils Valeinis, Dirk van den Berge, Rolf-Peter Mueller, Gloria Tridello, Laurence Collette and Andrew Bottomley

ABSTRACT

Purpose This phase III trial compared adjuvant whole-brain radiotherapy (WBRT) with observation after either surgery or radiosurgery of a limited number of brain metastases in patients with stable solid tumors. Here, we report the health-related quality-of-life (HRQOL) results.

Patients and Methods HRQOL was a secondary end point in the trial. HRQOL was assessed at baseline, at 8 weeks, and then every 3 months for 3 years with the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire C30 and Brain Cancer Module. The following six primary HRQOL scales were considered: global health status; physical, cognitive, role, and emotional functioning; and fatigue. Statistical significance required $P \leq .05$, and clinical relevance required a $\geq 10$-point difference.

Results Compliance was 88.3% at baseline and dropped to 45.0% at 1 year; thus, only the first year was analyzed. Overall, patients in the observation only arm reported better HRQOL scores than did patients who received WBRT. The differences were statistically significant and clinically relevant mostly during the early follow-up period (for global health status at 9 months, physical functioning at 8 weeks, cognitive functioning at 12 months, and fatigue at 8 weeks). Exploratory analysis of all other HRQOL scales suggested worse scores for the WBRT group, but none was clinically relevant.

Conclusion This study shows that adjuvant WBRT after surgery or radiosurgery of a limited number of brain metastases from solid tumors may negatively impact some aspects of HRQOL, even if these effects are transitory. Consequently, observation with close monitoring with magnetic resonance imaging (as done in the EORTC trial) is not detrimental for HRQOL.
INTRODUCTION

The role of whole-brain radiotherapy (WBRT) after either surgical resection or radiosurgery of brain metastases has been debated for many years.\(^1\) Two randomized trials found that the omission of WBRT in patients with newly diagnosed brain metastases either after complete surgical resection\(^3\) or stereotactic radiosurgery\(^4\) results in a significantly worse local and distant control in the brain on magnetic resonance imaging (MRI), but this does not affect overall and functionally independent survival. Our recently published European Organisation for Research and Treatment of Cancer (EORTC) 22952-26001 trial\(^5\) confirmed the results of these two studies in a larger patient population; the median time to deterioration to WHO performance status more than 2 and the overall survival were similar in the WBRT and observation arms, whereas a modest advantage in terms of progression-free survival was observed for WBRT.

Because the majority of patients with brain metastases still cannot be cured and treatments are directed toward the palliation of symptoms, maintenance of health-related quality of life (HRQOL) is a primary objective.\(^6\) Therefore, an evaluation of the influence of WBRT on HRQOL and patient-reported symptoms was undertaken as part of this study, and the results are reported here.

PATIENTS AND METHODS

Study design and treatment

This was an international multicenter phase III study (Fig 1) with the following two recruitment strata: radiosurgery patients were randomly assigned within 2 weeks before radiosurgery to either adjuvant WBRT or observation (arms 1 and 2, respectively) and surgery patients were randomly assigned within 4 weeks after surgery to either WBRT or observation (arms 3 and 4, respectively). The trial was conducted in accordance with the Declaration of Helsinki. All patients provided written informed consent before random assignment. The EORTC 22952-26001 study is registered with ClinicalTrials.gov (identifier: NCT00002899).

Procedures for HRQOL Data Collection

The two HRQOL measures were the EORTC Quality of Life Questionnaire C30 (EORTC QLQ-C30; version 3)\(^7\) and the EORTC QLQ Brain Cancer Module (EORTC QLQ-BN20).\(^8,9\) The EORTC QLQ-C30\(^10\) comprises five functioning scales (physical, role, emotional, cognitive, and social), three symptom scales (fatigue, nausea/vomiting, and pain), six single-item scales (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial impact), and the global health scale. The items on the measures were scaled and scored using recommended EORTC procedures.\(^11\) The EORTC QLQ-BN20 was developed specifically for patients with brain cancer\(^8,12\) and consists of 20 questions measuring four scales (visual disorder, motor dysfunction, communication deficit, and future uncertainty) and seven single items (headaches, seizures, drowsiness, hair loss, itchy skin, weakness of legs, and bladder control).\(^13\) Assessments were performed at random assignment, at 8 weeks after start of local treatment, either surgery or radiosurgery, and every 3 months thereafter for 3 years or until the patient's WHO performance status exceeded 2. Guidelines for administering questionnaires were provided, ensuring standardization of HRQOL data by all personnel.\(^14\)
Statistical analyses

HRQOL was a secondary study end point, and the sample size was based on the primary end point (ie, time to deterioration of WHO performance status to > 2). All analyses were performed using SAS (version 9.2; SAS Institute, Cary, NC), according to the intent-to-treat principle. The results of this study are presented in accordance with EORTC criteria for reporting HRQOL.15

Received HRQOL forms were considered as valid if, according to their completion date, they fell within the acceptable time window for their assessment. The definitions of the assessment time windows are as follows. Baseline assessment was between 4 weeks before random assignment, but not before any prior surgery, and the day of random assignment, but no later than the start of the local treatment. The 8-week assessment was due between 2 and 6 weeks after start of local treatment. The 3-month assessments were due at 3, 6, 9, and 12 months after the 8-week assessment, falling no earlier or later than 6 weeks from the intended assessment time. If two or more HRQOL forms were received within the same time window, the form closest to the scheduled assessment time was kept.

At the time of protocol writing, no studies had provided clear guidance on the directionality of HRQOL scales for the EORTC QLQC-30 and the EORTC QLQ-BN20 in this population. Therefore, no HRQOL hypothesis for individual scales was suggested. Before the analysis, an HRQOL analysis plan was devised. To reduce the problem of multiplicity, the primary HRQOL
analysis was performed on selected HRQOL domains, namely those that were expected to be affected by the WBRT and to be considered of relevance to the patients' daily life. The following six scales were selected: global health status; physical, cognitive, role, and emotional functioning; and fatigue. The remaining HRQOL variables were examined on an exploratory basis. The overall HRQOL hypothesis was that WBRT, by providing an increase in intracerebral tumor control and prolonging survival with WHO performance status $\leq 2$, would lead to both short- and long-term improved global quality of life. No scale-specific hypotheses were generated because no consensus on the expected differences could be found as a result of the lack of reliable historical data.

A linear mixed model approach, including interaction between treatment and time, stratum (radiosurgery v surgery), localization of the primary tumor (lung v other), and presence of macroscopic tumor outside the brain as fixed effects, estimated the HRQOL differences with a flexible covariance structure. For each scale, all HRQOL scores (including baseline) were used as dependent outcomes in the model. Baseline scores were not added as a covariate to the model. Score estimates, SEs, the associated CIs, and resulting tests were obtained from the model, including a general overall postbaseline test formed by testing a specific hypothesis for the entire population from which the random effects are sampled. The hypothesis is that the differences between the two treatment arms for all postbaseline time points (ie, 8 weeks and 3, 6, 9, and 12 months) equal 0 via an overall F test statistic. Tests for treatment difference at each time point were obtained by contrasting the treatment-time interaction covariates at a given time point between the two arms via an F test.

A standard method of interpretation was used for the HRQOL using the EORTC scores according to the minimal important difference approach. Differences of at least 10 points (on a 0 to 100 scale) were classified as the minimum clinically meaningful change in an HRQOL parameter. The level of statistical significance was fixed at $P = .05$. Because missing data are a problem in most HRQOL studies, sensitivity analyses were performed investigating the informative dropout by graphical evaluation and which variables impact on the compliance via linear regression. For the primary HRQOL scales, explicit regression imputation was applied on the intent-to-treat population where imputed values were predicted from a regression model including factors related to a missingness mechanism applied to the observed data. The percentage of patients who experienced at least a clinically relevant worsening in the postbaseline time points was calculated and compared in the two treatment groups. The group of patients who had baseline and follow-up HRQOL data provided 90% power to detect a 10-point shift on the global HRQOL scale, using a two-sided test at the 5% significance level. This post hoc estimation of the power was based on the average observed standard deviation at first follow-up visit of the global health scale score (standard deviation, 22).

**RESULTS**

From November 1996 to November 2007, 359 patients were randomly assigned (Fig 1), 180 patients in the WBRT arm (99 in the radiosurgery stratum and 81 in the surgery stratum) and 179 in the observational arm (100 in the radiosurgery stratum and 79 in the surgery stratum). Baseline characteristics were compared between patients with ($n = 341$) and without ($n = 18$) valid HRQOL
forms. A statistical difference was found for neurologic status (P < .001); the median time to WHO performance status deterioration was shorter for patients with no HRQOL form at all than for patients with at least one HRQOL form.

**Compliance With HRQOL Measures**

Compliance was 88.3% at baseline, 62% at week 8, 60% at 3 months, and still greater than 50% up to 9 months. Compliance dropped to less than 50% at 12 months (45%, 65 patients). Because of the small number of patients with data still available at 15 months (n = 56), data were analyzed up to 12 months. At baseline, patients who had surgery alone showed lower compliance than did patients who had radiosurgery (80.0% v 95.0%, respectively; P < .001). χ² tests for compliance differences between the two treatments arms (controlling for the stratum radiosurgery v surgery) revealed no significant differences at baseline and at any follow-up time points (Table 1).

<table>
<thead>
<tr>
<th>Assessment Time</th>
<th>No. of Forms Received</th>
<th>No. of Forms Expected</th>
<th>Compliance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>317</td>
<td>359</td>
<td>88.3</td>
</tr>
<tr>
<td>WBRT</td>
<td>162</td>
<td>180</td>
<td>90.0</td>
</tr>
<tr>
<td>OBS</td>
<td>155</td>
<td>179</td>
<td>86.6</td>
</tr>
<tr>
<td>8 weeks</td>
<td>206</td>
<td>333</td>
<td>61.9</td>
</tr>
<tr>
<td>WBRT</td>
<td>105</td>
<td>169</td>
<td>62.1</td>
</tr>
<tr>
<td>OBS</td>
<td>101</td>
<td>164</td>
<td>61.6</td>
</tr>
<tr>
<td>3 months</td>
<td>156</td>
<td>262</td>
<td>59.5</td>
</tr>
<tr>
<td>WBRT</td>
<td>81</td>
<td>133</td>
<td>60.9</td>
</tr>
<tr>
<td>OBS</td>
<td>75</td>
<td>129</td>
<td>58.1</td>
</tr>
<tr>
<td>6 months</td>
<td>107</td>
<td>210</td>
<td>51.0</td>
</tr>
<tr>
<td>WBRT</td>
<td>53</td>
<td>105</td>
<td>50.5</td>
</tr>
<tr>
<td>OBS</td>
<td>54</td>
<td>105</td>
<td>51.4</td>
</tr>
<tr>
<td>9 months</td>
<td>88</td>
<td>170</td>
<td>51.8</td>
</tr>
<tr>
<td>WBRT</td>
<td>45</td>
<td>87</td>
<td>51.7</td>
</tr>
<tr>
<td>OBS</td>
<td>43</td>
<td>83</td>
<td>51.8</td>
</tr>
<tr>
<td>12 months</td>
<td>65</td>
<td>144</td>
<td>45.1</td>
</tr>
<tr>
<td>WBRT</td>
<td>29</td>
<td>73</td>
<td>39.7</td>
</tr>
<tr>
<td>OBS</td>
<td>36</td>
<td>71</td>
<td>50.7</td>
</tr>
</tbody>
</table>

**Table 1. Compliance With Health-Related Quality-of-Life Assessments**

NOTE. Definitions of the assessment times are provided in Patients and Methods. Abbreviations: OBS, observation; WBRT, whole-brain radiotherapy.

**Primary HRQOL Scales: Differences Between Arms and Changes Over Time**

The global HRQOL scores at baseline were similar in both treatment groups. However, a statistically significant and clinically meaningful difference in global HRQOL mean scores was detected at 9 months in favor of patients who had observation alone (mean, 63.2; SE, 3.2 for observation v mean, 52.2; SE, 3.2 for WBRT; P = .0148), whereas no differences were found at any other time points or considering the post–random assignment overall difference (Table 2; Fig 2).
Patients in the observation only group had better mean scores in physical, role, and cognitive functioning (Table 3; Fig 2). The mean difference was statistically significant for physical functioning at 8 weeks (mean, 52.3; SE, 2.9 for observation v mean, 41.6; SE 2.8 for WBRT; P = .0073), role functioning at 8 weeks (mean, 66.9; SE, 3.1 for observation v mean, 58.3; SE, 3.0 for WBRT; P = .0491), and cognitive functioning at 8 weeks (mean, 81.2; SE, 2.3 for observation v mean, 73.9; SE, 2.3 for WBRT; P = .0263) and at 12 months (mean, 80.4; SE, 3.7 for observation v mean, 69.7; SE, 4.0 for WBRT; P = .0486). No postbaseline differences were found for emotional functioning. For fatigue, the mean difference was statistically significant at 8 weeks (mean, 50.8; SE, 2.5 for WBRT v mean, 38.9; SE, 2.6 for observation; P < .001) and at 3 months (mean, 49.5; SE, 2.8 for WBRT v mean, 41.3; SE, 2.9 for observation; P = .0438).
Other HRQOL Scales

Table 3 also contains the results of the exploratory analysis of the nonpreselected remaining HRQOL scales, which showed statistically significant worse scores for bladder control, communication deficit, drowsiness, hair loss, motor dysfunction, weakness in the legs, appetite loss, constipation, nausea/vomiting, pain, and social functioning (considering the overall postbaseline scores) in patients who underwent WBRT compared with observation.

Missingness Mechanism and Sensitivity Analysis

For each of the preselected HRQOL scales (global HRQOL; physical, cognitive, emotional, and role functioning; and fatigue), the evolution of the mean scores just before dropout was graphically investigated to check the missingness mechanism. For role functioning, fatigue, and, more notably, physical functioning, a decrease in scores just before the dropout time was observed, indicating an informative selection effect. Therefore, a sensitivity analysis with explicit regression imputed scores was performed.
Scores split by stratum (radiosurgery and surgery) were analyzed as a sensitivity analysis. In general, these results confirmed the primary analyses, but for physical functioning, differences in trend between the two strata were found. Because of a substantial proportion of missing data, explicit regression imputation was applied for the primary HRQOL scales selected a priori based on a generalized linear model including factors related to the missingness mechanism, such as sex, treatment (WBRT v observation), stratum (radiosurgery v surgery), and the interaction between the performance status (2 v 0 to 1) and the expected number of assessments (from one to six assessments). A linear mixed model was then performed on the augmented data set; these results remained similar. Tests adjusted for baseline were also performed, and the observed treatment differences tended to decrease. Finally, the percentage of patients who experienced at least a clinical relevant worsening in the postbaseline time points was compared in the two groups, and no difference was noticed between the two treatments.

### DISCUSSION

HRQOL is a well-established end point for treatment comparisons in many cancer types, particularly in advanced stages.\(^{18}\) HRQOL instruments, in particular the EORTC QLQ-BN20 module, originally developed for primary brain tumors, seem appropriate to cover the whole range of symptoms and HRQOL issues caused by brain metastases and their treatments.

HRQOL has been analyzed in trials investigating the role of WBRT with or without radiotherapy sensitizers or chemotherapy for patients with multiple brain metastases from solid tumors.\(^{19-21}\) No
data are available so far from phase III trials for patients with a limited number of brain metastases (mostly single lesions) undergoing surgical resection or radiosurgery, who can either receive adjuvant WBRT or be observed with MRI and delay treatment until tumor progression. In this regard, both the American\(^3\) and Japanese\(^4\) trials used survival and local tumor control end points only. Our trial used HRQOL as a secondary end point and thus provides this unique patient perspective not provided by the traditional outcome measures. To our knowledge, this trial is the first randomized controlled trial reporting a prospective analysis and comparison of HRQOL and patient-reported symptoms in patients with brain metastases, who were either treated with WBRT or observed with MRI after local treatments.

Overall, patients in observation only arm reported better HRQOL scores than did patients who received WBRT. Most scores, which differed significantly during the first time points, had a tendency to recover. We noted a significant and clinically relevant (≥ 10 points) difference in global health status in favor of the observation arm at 9 months (P = .0148). Furthermore, we observed a significant and clinically relevant difference for physical functioning at 8 weeks (10.7 points) and cognitive functioning at 12 months (10.7 points) in favor of observation alone. Patients receiving WBRT reported more fatigue at 8 weeks (difference of 11.9 points); it is possible that fatigue may adversely affect cognitive functioning.\(^22\) Moreover, WBRT had an adverse impact on hair loss, appetite loss, constipation, nausea/vomiting, drowsiness, and social functioning. Similarly, an EORTC phase III trial on prophylactic cranial irradiation (PCI) in extensive small-cell lung cancer reported a short-term (within 3 months) negative impact of PCI on selected HRQOL scales.\(^22\) The largest mean difference between the two arms (PCI v control) was observed for fatigue and hair loss, whereas the negative impact of PCI on the global health status and functioning scales was minimal.

Our trial cohort consists largely of patients with favorable prognostic factors (high performance status, single lesions, and stable/absent systemic cancer)\(^23,24\) and who thus are candidates for longer survival; in this population, the aim of adjuvant WBRT is to prevent neurologic/neurocognitive and HRQOL deterioration, which is associated with tumor progression.\(^25,26\) In the trial, the positive effect of WBRT in decreasing the rate of intracranial progression and modestly improving the progression-free survival\(^5\) did not translate into an advantage in terms of HRQOL. The following two reasons are possible: patients who were observed with MRI only had early detection of asymptomatic brain relapses, and salvage treatments were able to prevent a deterioration of HRQOL. Also, patients who received WBRT experienced more systemic progresses (46% v 33.5% for observation), often requiring more intensive treatments that could have a negative impact on HRQOL.

Our study has a number of limitations. The major challenge is the level of compliance, and although compliance was high at baseline, it decreased over subsequent assessments and dropped to less than 50% at 12 months (45%). Overall, it remained within acceptable limits for studies in advanced disease\(^27\) and was either better than that reported in other studies on brain metastases\(^19,20\) or comparable with studies in primary brain tumors such as glioblastoma.\(^28\) The comparison of baseline characteristics between patients with and without valid HRQOL questionnaires showed a significant difference for neurologic status (P < .001); the median time to WHO performance status deterioration was shorter for patients with no HRQOL forms than for patients with at least one HRQOL form. This might be a result of the fact that WHO performance status deterioration was a
condition to withdraw from the HRQOL assessment schedule. Nevertheless, we are aware that missing data are a significant threat for the interpretation of patients' HRQOL, limiting the power of statistical analyses. The long-term impact of WBRT on HRQOL could not be assessed because the number of patients became too small to detect any reliable change in HRQOL after the first year.

Another limitation of this study is that, although we assessed cognitive functioning with the EORTC QLQ-C30, we did not assess cognitive function with cognitive test batteries, and it is known that the self-report of cognitive functioning and formal neurocognitive testing may be poorly correlated. In patients with brain metastases, neurocognitive deterioration can either precede and negatively influence patients' HRQOL\textsuperscript{21} or represent a subtle form of neurotoxicity from WBRT,\textsuperscript{1} especially in long-term survivors. There is still a lack of data regarding long-term cognitive changes with treatment (surgery, radiosurgery, WBRT, or chemotherapy). Aoyama et al\textsuperscript{29} serially assessed neurocognitive function using the Mini-Mental State Examination in the Japanese trial\textsuperscript{4} and observed some deterioration of neurocognitive function among long-term survivors who underwent WBRT. These data must be interpreted with caution, because the Mini-Mental State Examination has poor sensitivity in detecting neurocognitive problems in brain tumor patients\textsuperscript{30}; thus, the use of specific neurocognitive batteries of tests has now become mandatory in clinical trials.\textsuperscript{31} Two recent randomized trials have reported early negative changes with respect to neurocognitive function after WBRT. The first trial,\textsuperscript{32} which compared radiosurgery alone with radiosurgery plus WBRT in patients with brain metastases, reported that patients receiving the combined treatment were at greater risk of a significant decline in learning and memory function by 4 months. The second trial,\textsuperscript{33} which compared PCI with observation in patients with locally advanced non–small-cell lung cancer, reported a greater deterioration in some memory functions (ie, immediate and delayed recall on the Hopkins Verbal Learning Test) at 3 months. Overall, this latter trial seemed to parallel the results of the present trial; in fact, in both studies, the benefit of WBRT in terms of reduction of brain relapses was counterbalanced by a lack of overall survival advantage and a short-term negative impact on quality of life and cognitive functioning.

In summary, our study has shown that adjuvant WBRT after surgery or radiosurgery of a limited number of brain metastases from solid tumors may negatively impact some aspects of HRQOL, even if these effects are transitory. As a result, choosing observation with close monitoring with MRI (as done in the EORTC trial), instead of adjuvant WBRT, is not detrimental to HRQOL.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Riccardo Soffietti, Martin Kocher, Rolf-Dieter Kortmann, Dirk van den Berge, Rolf-Peter Mueller, Andrew Bottomley

Administrative support: Gloria Tridello, Laurence Collette, Andrew Bottomley

Provision of study materials or patients: Riccardo Soffietti, Ufuk M. Abacioglu, Salvador Villa, François Fauchon, Brigitta G. Baumert, Laura Fariselli, Tzahala Tzuk-Shina, Rolf-Dieter
Kortmann, Christian Carrie, Mohamed Ben Hassel, Mauri Kouri, Egils Valeinis, Dirk van den Berge, Rolf-Peter Mueller

Collection and assembly of data: Salvador Villa, Tzahala Tzuk-Shina, Christian Carrie, Mohamed Ben Hassel, Mauri Kouri, Egils Valeinis, Dirk van den Berge

Data analysis and interpretation: Riccardo Soffietti, Ufuk M. Abacioglu, Salvador Villa, François Fauchon, Brigitta G. Baumert, Laura Fariselli, Gloria Tridello, Laurence Collette, Andrew Bottomley

Manuscript writing: All authors

Final approval of manuscript: All authors

ACKNOWLEDGEMENT

We thank all European Organisation for Research and Treatment of Cancer (EORTC) investigators who participated in this study and the EORTC Headquarters staff who worked for this study. We thank all the patients for participating in this clinical trial to help us understand the best way to treat this cancer. We also thank Corneel Coens from the EORTC Headquarters for his methodologic support in the analysis of the quality-of-life data.

REFERENCES


