

OXFORD

JNCI: Natl Cancer Inst (2016) 108(5)-djv391

doi:10.1093/jnci/djv391

First published online December 28, 2015

Review

REVIEW

The Added Value of Analyzing Pooled Health-Related Quality of Life Data: A Review of the EORTC PROBE Initiative

Efstathios Zikos, Corneel Coens, Chantal Quinten, Divine E. Ediebah, Francesca Martinelli, Irina Ghislain, Madeleine T. King, Carolyn Gotay, Jolie Ringash, Galina Velikova, Bryce B. Reeve, Eva Greimel, Charles S. Cleeland, Henning Flechtner, Martin J. B. Taphoorn, Joachim Weis, Joseph Schmucker-von Koch, Mirjam A. G. Sprangers, Andrew Bottomley; on behalf of EORTC PROBE

Affiliations of authors: European Organisation for Research and Treatment of Cancer Headquarters, Brussels, Belgium (EZ, CC, DEE, FM, IG, AB); European Centre for Disease Prevention and Control, Surveillance and Response Support, Epidemiological Methods Unit, Stockholm, Sweden (CQ); School of Psychology and Sydney Medical School, University of Sydney, New South Wales, Australia (MTK); School of Population and Public Health, University of British Columbia, Vancouver, Canada (CG); The Princess Margaret Cancer Centre, University of Toronto, Toronto, Canada (JR); Leeds Institute of Cancer and Pathology, University of Leeds, St. James's Hospital, Leeds, UK (GV); Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC (BBR); Obstetrics and Gynecology, Medical University Graz, Graz, Austria (EG); Department of Symptom Research, Division of Internal Medicine, The University of Texas Anderson Cancer Center, Houston, TX (CSC); Clinic for Child and Adolescent Psychiatry and Psychotherapy, University of Magdeburg, Magdeburg, Germany (HF); VU University Medical Center / Medical Center Haaglanden, Amsterdam/The Hague, the Netherlands (MJB/T); Tumor Biology Center, University of Freiburg, Freiburg, Germany (JW); Philosophical Faculty, University of Regensburg, Regensburg, Germany (JSvK); The Academic Medical Center, Amsterdam, the Netherlands (MAGS).

Correspondence to: Efstathios Zikos, MS, MA, Quality of Life Department, EORTC Headquarters, Avenue E. Mounier 83, 1200, Brussels, Belgium (e-mail: efstathios.zikos@eortc.be).

Abstract

Background: The European Organisation for Research and Treatment of Cancer (EORTC) Patient-Reported Outcomes and Behavioural Evidence (PROBE) initiative was established to investigate critical topics to better understand health-related quality of life (HRQOL) of cancer patients and to educate clinicians, policy makers, and healthcare providers.

Methods: The aim of this paper is to review the major research outcomes of the pooled analysis of HRQOL data along with the clinical data. We identified 30 pooled EORTC randomized controlled trials (RCTs), 18 NCIC-Clinical Trials Group RCTs, and two German Ovarian Cancer Study Group RCTs, all using the EORTC QLQ-C30. All statistical tests were two-sided.

Results: Evidence was found that HRQOL data can offer prognostic information beyond clinical measures and improve prognostic accuracy in cancer RCTs (by 5.9%-8.3%). Moreover, models that considered both patient- and clinician-reported scores gained more prognostic overall survival accuracy for fatigue ($P < .001$), vomiting ($P = .01$), nausea ($P < .001$), and constipation ($P = .01$). Greater understanding of the association between symptom and/or functioning scales was developed by identifying physical, psychological, and gastrointestinal clusters. Additionally, minimally important differences in interpreting HRQOL changes for improvement and deterioration were found to vary across different patient populations and disease stages. Finally, HRQOL scores are statistically significantly affected by deviations from the intended time point at which the questionnaire is completed.

Received: July 10, 2015; Revised: October 21, 2015; Accepted: November 20, 2015

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

Conclusions: The use of existing pooled data shows that it is possible to learn about general aspects of cancer HRQOL and methodology. Our work shows that setting up international pooled datasets holds great promise for understanding patients' unmet psychosocial needs and calls for additional empirical investigation to improve clinical care and understand cancer through retrospective HRQOL analyses.

The demand for symptom relief, reduced side effects of medical strategies, and improved patient satisfaction with care led to the increased collection and analysis of patient-reported health-related quality of life (HRQOL) data. Such data help determine the effectiveness of clinical interventions from the patient perspective. Yet several methodological issues remain in the assessment, analysis, and interpretation of HRQOL data from clinical trials. We pooled the data of multiple closed randomized controlled trials (RCTs) in order to address a number of methodological questions of relevance to the field that were not addressed by the original trials. An overview of the results of the European Organisation for Research and Treatment of Cancer (EORTC) Patient-Reported Outcomes and Behavioural Evidence (PROBE) project (1) initiative are reported here.

The PROBE project was launched in January 2008 as an initiative of the EORTC Quality of Life Department (QLD). QLD has provided support and expertise toward the inclusion of HRQOL endpoints in EORTC RCTs since 1993, following the development and publication of the EORTC Quality of Life Core Questionnaire (QLQ-C30) (2). The QLQ-C30 is comprised of 30 questions measuring fifteen HRQOL parameters: five functioning scales (physical, role, cognitive, emotional, and social), three symptom scales (pain, fatigue, nausea/vomiting), six single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial problems), and global health status/quality of life. All scales and single item measures range in score from 0 to 100. Higher scores represent better health for the functional scales and the global health status/quality of life scale while for the symptom scales and the symptom items higher scores represent worse health. The QLQ-C30, which is applicable to the general cancer population, has become the most widely used HRQOL questionnaire in cancer RCTs during the last two decades (3–7).

This review arises from the activities of the EORTC PROBE initiative, which aimed to develop a user-ready HRQOL database of scores collected on the QLQ-C30 and to use it to investigate general research questions common across different cancer types. Collecting data from closed international oncology RCTs

that informed clinical practice over the years (8) has been challenging.

HRQOL data are collected and analyzed trial by trial and published either together with clinical data or separately to measure the impact of cancer and its treatment on patients' HRQOL. This leads inevitably to a fragmented body of evidence, often with inconsistent methods of data collection, analysis, and reporting. Construction of a meta-dataset, which pools data from different cancer clinical trials across different patient populations (eg, lung, breast, melanoma) or disease stages (eg, primary vs advanced) can give a more comprehensive view of HRQOL in oncology. Moreover, it allows the application of analytical techniques demanding large sample sizes in a field where data collection is often very restricted and expensive. PROBE undertook the challenge to merge international RCTs and make meaningful analyses of these data with the support of the PROBE international advisory board (composed of medical, clinical, statistical, psychological, and other experts across different fields), which scrutinized the analyses and their interpretation and advised on the overall management of the research initiative.

PROBE had a specific mandate to focus on the practical application of HRQOL in clinical trials research. The research objectives identified by the PROBE team and reviewed in this article were therefore to assess the prognostic value of HRQOL for survival, to compare clinicians' and patients' HRQOL assessment for their prognostic value for overall survival (OS), to explore whether QLQ-C30 domains cluster, to investigate minimal important differences (MIDs) for interpreting HRQOL scores from the QLQ-C30, and to examine the effect of completion time windows on HRQOL outcomes (Figure 1).

Methods

The PROBE Dataset

Over the last 25 years, EORTC used HRQOL as a primary or secondary endpoint in over 150 clinical trials. Many of these

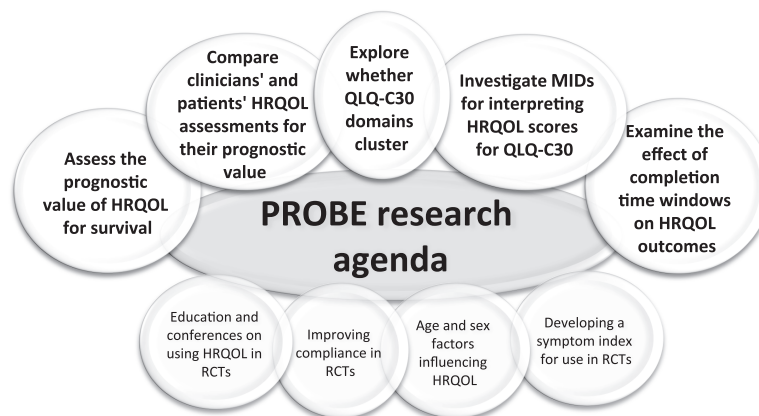


Figure 1. The Patient-Reported Outcomes and Behavioural Evidence research agenda. HRQOL = health-related quality of life; MIDs = minimally important differences; PROBE = Patient-Reported Outcomes and Behavioural Evidence; QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Core Questionnaire; RCTs = randomized controlled trials.

trials were either open to patient recruitment or the primary endpoint results were not yet published during the selection of trials (either because of early termination or primary endpoint being analyzed) or did not have good HRQOL compliance data, making it unfeasible to use those RCT data. Finally, thirty of these RCTs with QLQ-C30 data, good patient compliance, and where EORTC QLD obtained the permission for data use were selected. Those RCTs were conducted from 1986 to 2004, and included 10 874 patients from 11 different cancer sites. The data from these RCTs were extracted from the EORTC database and merged to form the PROBE database. In addition, other research organizations were invited to share their RCT data that used the QLQ-C30 and cancer-related modules. Permission for eighteen RCTs from the NCIC-Clinical Trials Group (NCIC-CTG), including 4635 patients' data from 15 different cancer sites, and two ovarian cancer RCTs with 1731 patients' data from the German Ovarian Cancer Study Group (AGO) were obtained and shared via external software platforms (Figure 2).

All selected international RCTs included HRQOL as a secondary outcome collected via the QLQ-C30. Standardized socio-demographic and clinical data (eg, age, sex, World Health Organization [WHO] performance status, and stage of disease) were also included. The PROBE database includes data from a total of 17 239 individual patients with completed QLQ-C30, with or without its supplementary modules, and clinical and survival outcomes.

Statistical Analysis

The different analysis techniques and statistical methods used in the reviewed research projects are fully reported within the references of the previously conducted analyses (9–17). All statistical tests were two-sided, and a *P* value of less than .05 was considered statistically significant.

Results

HRQOL Adds Prognostic Value Beyond Clinical Information

One of the key questions cancer patients ask their clinicians when diagnosed with cancer is ‘How long will I live?’ Research into prognostic indicators of survival is an important topic in oncology, and data from many single RCTs have identified various factors. A previously published review of HRQOL data collected in cancer RCTs (5) indicated that HRQOL data are prognostic of survival above and beyond clinical predictors. Among the 39 reviewed RCTs by Gotay et al. (5), four functioning (physical, role, social, emotional) and three symptom scales (pain, fatigue, nausea, and vomiting), along with four single items (appetite loss, dyspnea, insomnia, constipation) and global health status/quality of life, were reported to be prognostic of survival. Four PROBE analyses provided further evidence that HRQOL can improve the prognostic value in cancer clinical trials (9–12).

First, Quinten et al. (9) examined the prognostic significance of socio-demographic (age and sex), clinical variables (WHO performance status, distant metastasis, and cancer site) and the 15 QLQ-C30 scales and individual items across different disease sites. In the Cox multivariable model including the aforementioned variables, the HRQOL domains physical functioning ($P < .0001$), pain ($P < .0001$) and appetite loss ($P < .0001$) provided statistically significant prognostic information in addition to age ($P = .0001$), sex ($P < .0001$), and distant metastases ($P < .0001$); meanwhile, the WHO performance status did not add prognostic information ($P = .11$). The three HRQOL domains improved the estimation of survival by 5.9% relative to socio-demographic and clinical characteristics alone. A second survival analysis (10) in a subgroup of 2410 metastatic cancer patients revealed that physical functioning, pain, and appetite loss, along with the

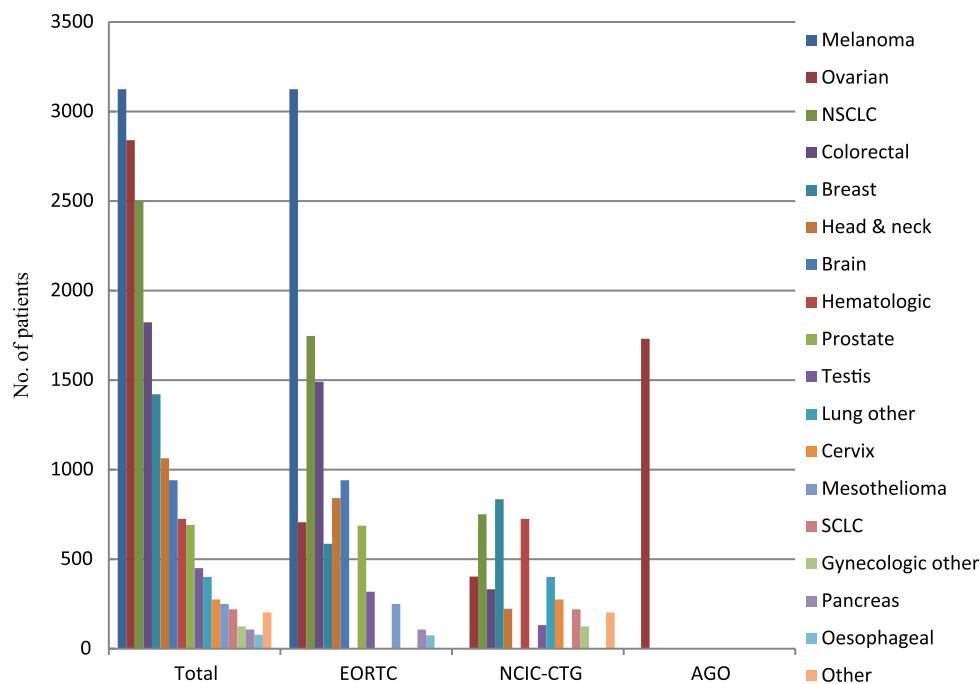


Figure 2. The Patient-Reported Outcomes and Behavioural Evidence dataset subgrouped by cancer sites. AGO = German Ovarian Cancer Study Group; EORTC = European Organisation for Research and Treatment of Cancer; NCIC-CTG = NCIC-Clinical Trials Group; NSCLC = non-small cell lung cancer; PROBE = Patient-Reported Outcomes and Behavioural Evidence; SCLC = small cell lung cancer.

variables age, sex, and WHO performance status, increased prognostic accuracy by 8.3% compared with clinical variables alone, thereby demonstrating that the prognostic value of HRQOL may differ across clinical subgroups.

The question then arose of whether or not different HRQOL domains are prognostic for survival for different cancer sites. By using a dataset of 7417 patients who completed the QLQ-C30 before randomization, Quinten et al. (11) found that at least one HRQOL domain provided prognostic information for each cancer site, alongside clinical and socio-demographic variables (Table 1), although which domain provided the greatest prognostic power differed by cancer type.

A universal HRQOL domain with valid prognostic impact across cancer sites could not be identified. Physical functioning and nausea/vomiting were found to provide unique prognostic information in several but not all cancer sites.

By using longitudinal HRQOL data from a single advanced non-small cell lung cancer trial (n = 391), we investigated whether changes in HRQOL scores over time during chemotherapy treatment can be prognostic of survival in addition to clinical characteristics. It was found that a 10-point improvement in pain (from baseline to end of cycle 1) or social function (from baseline to end of cycle 2) was associated with a lower risk of death (12). Meanwhile pain, physical function, and dysphagia (single item of lung cancer-specific module QLQ-LC13 assessing treatment-related side effects [18]) were important baseline prognostic factors.

These studies provided further evidence that HRQOL can provide prognostic information beyond clinical measures, improve prognostic accuracy in cancer clinical trials, and yield informative factors to stratify and monitor patients for supportive interventions but that the actual prognostic HRQOL domains may vary across disease sites (9–12).

Table 1. Multivariable Cox regression analyses of survival (HRs of survival [95% CIs for socio-demographic, clinical, and HRQOL scales across the 11 cancer sites])*

Cancer site	Socio-demographic/clinical scales		HRQOL scales		
			HR (95%CI) P		
Colorectal	Age (≤60 vs >60 y)	Sex (male vs female)	Nausea and vomiting	Physical functioning 0.93 (0.90 to 0.99)	Appetite loss 1.07 (1.04 to 1.11)
	1.27 (1.07 to 1.51) .0011	0.81 (0.68 to 0.98) .0094	1.07 (1.04 to 1.11) <.0001	<.0001	<.0001
Lung	WHO PS (good vs poor)	Sex (male vs female)	Pain	Physical functioning 0.94 (0.90 to 0.98)	
	1.64 (1.18 to 2.82) .0005	0.76 (0.62 to 0.91) .0008	1.08 (1.05 to 1.11) <.0001	.0006	
Prostate	WHO PS (good vs poor)	-	Role functioning 0.96 (0.94 to 0.98)	Appetite loss 1.07 (1.03 to 1.10)	
	1.37 (1.15 to 1.63) <.0001		.006	<.0001	
Brain	WHO PS (good vs poor)	Age (≤60 vs >60 y)	-	Cognitive functioning 0.95 (0.90 to 0.97)	
	1.68 (1.36 to 2.10) <.0001	1.55 (1.28 to 1.88) <.0001		<.0001	
Breast	WHO PS (good vs poor)	Distant metastasis (no vs yes)	-	Nausea and vomiting 1.17 (1.05 to 1.29)	
	4.12 (1.78 to 9.52) <.0001	21.45 (1.78 to 258.43) <.0001		.0011	
Melanoma	WHO PS (good vs poor)	Sex (male vs female)	-	Dyspnea 1.06 (1.03 to 1.10)	
	1.71 (1.08 to 2.70) .0064	0.61 (0.50 to 0.72) <.0001		<.0001	
Head & neck	-	Sex (male vs female)	-	Nausea and vomiting 1.14 (1.01 to 1.27)	
		0.38 (0.21 to 0.71) .0003		.0097	
Esophageal	-	-	-	Physical functioning 0.88 (0.80 to 0.96)	
				.0072	
Ovarian	-	-	-	Nausea and vomiting 1.2 (1.10 to 1.30)	
				<.0001	
Testicular	-	-	-	Role functioning 0.81 (0.67 to 0.95)	
				.0144	
Pancreas	-	-	-	Global health status/quality of life scale 0.83 (0.71 to 0.95)	
				.0073	

* CI = confidence interval; HR = hazard ratio; HRQOL = health-related quality of life; WHO PS = World Health Organization performance status.

Patients' Self-Reports Along With Clinicians' Scores Improve Survival Prediction

Typically, to estimate overall survival, clinician evaluation of symptoms is incorporated into a model of prognosis. When considering treatment options, a reliable survival prognosis is a valuable tool to make an informed decision. The clinician scoring of symptoms has conventionally been used. However, the weak agreement between clinician and patient reporting of symptoms is notable (19). In this PROBE analysis, the relative information gained in estimating survival when including baseline information of patient-reported symptoms was compared with that reported by clinicians (assessed using Common Terminology Criteria for Adverse Events) across various disease sites, stages, and treatments. We found that for the six symptoms assessed at baseline—pain, fatigue, vomiting, nausea, diarrhea, and constipation—the models that considered both patient and clinician scores gained statistically significantly more prognostic OS accuracy, namely fatigue ($P < .001$), vomiting ($P = .01$), nausea ($P < .001$), and constipation ($P = .01$), than models that considered clinician scores alone (13). The results of this retrospective PROBE analysis were acknowledged by the American Society of Clinical Oncology's (ASCO) Annual Report on Progress Against Cancer as one among those with the greatest potential impact on patients' lives (20), suggesting that adding patient-reported symptom scores to the traditional physician-based scoring system may result in a more accurate prognosis of survival.

Examining HRQOL Domain Clusters

Understanding the grouping between symptom and/or functioning domains may aid clinicians in managing the symptom burden experienced by patients and may help policy makers develop psychosocial support plans. We attempted to identify how HRQOL domains cluster and which HRQOL indicators are linked to patients' perception of overall quality of life (14). The results revealed physical (physical and role functioning, fatigue, pain), psychological (emotional and cognitive functioning, insomnia), and gastrointestinal clusters (nausea/vomiting, appetite loss) emerging from the overall dataset. Each cluster had high to moderate internal consistency ($\alpha = .84$, $.64$, and $.67$, respectively), indicating that the included scales are associated. The same clusters were found in subgroups defined according to socio-demographic and clinical characteristics while some differences emerged among cancer sites. The global health status/quality of life scale was found to be part of the physical cluster in the overall dataset. This result was consistent across different levels of disease severity, but divergent results were seen across some cancer sites. These findings suggest that HRQOL domains are interrelated and form clusters; however, clusters vary by disease. Identifying the mechanisms that define the relationships between HRQOL domains is important for appropriate problem management and the identification of populations that could benefit from receiving tailored psychosocial support and/or improved supportive care interventions.

Providing More Evidence-Based Data on Minimally Important Differences

MID refers to the smallest change or difference between HRQOL scores that is considered to be clinically relevant. This is an important notion with many implications, as it informs

clinicians, patients, regulators, and clinical trialists as to which changes in HRQOL scores are important. For example, MIDs may be used to assess the value of a health care intervention or to compare treatments, to make adjustments in health care policies, or to inform a clinician's decision to apply an intervention in a given situation (21). MIDs may also be useful in determining sample sizes in designing future RCTs. The methods commonly used to calculate MID are anchor-based or distribution-based (22–25). An example of the most widely used anchor-based approach was provided in the first estimation of MIDs for the QLQ-C30, which used the subjective significance questionnaire (21) to link changes in QLQ-C30 scores to patients' ratings of subjectively meaningful changes.

A PROBE analysis of two closed non-small cell lung cancer (NSCLC) RCTs (26–27) demonstrated that MID estimates for improvement appeared larger than those for deterioration in QLQ-C30 (15). The WHO performance status and weight change were used as clinical anchors. When anchoring with WHO performance status, the MID estimates for improvement or deterioration respectively were: physical functioning: (9, 4), role functioning: (14, 5), social functioning: (5, 7), global health status/quality of life scale: (9, 4), fatigue: (14, 6), and pain: (16, 3). Statistically significant differences ($P < .05$) in HRQOL across groups defined by these anchors were noted for all scales except for social functioning. The results suggest that in patients with advanced NSCLC undergoing treatment MID may depend upon whether the patients' WHO performance statuses are improving or worsening, but the results are not definitive. Additional MID analysis was carried out using HRQOL data from two EORTC high-grade glioma brain cancer RCTs (28–29). The WHO performance status and the mini-mental state examination were used as clinical anchors to determine the MID in HRQOL change scores in the QLQ-C30 and the brain module, the QLQ-BN20. Based on WHO performance status, our findings provided the following estimates of the MID for improvement and deterioration respectively: physical: (6, 9), role: (14, 12), and cognitive functioning: (8, 8), and global health status/quality of life scale: (7, 4), fatigue: (12, 9), and motor dysfunction: (4, 5). The results suggested that MID estimates for improvement and for deterioration vary across the selected HRQOL scales; in these brain cancer studies, there was no clear indication that the MID for improvement was systematically larger or smaller than the MID for deterioration (16).

The above findings suggest that although current guidelines are helpful for general interpretation more research is needed to estimate and compare MIDs on the QLQ-C30 across various patient populations, eg, across different cancer sites (melanoma, ovarian, etc.) as well as across stages of disease. This is consistent with the recommendations of Revicki et al. (30) that the MID is not an immutable characteristic but may vary by population and context.

HRQOL Completion Time Windows: Do They Matter?

A key aspect in the design and analysis of HRQOL data is the timing of the assessments. In an appropriately designed clinical trial, the protocol will state exactly when the HRQOL assessments are scheduled, eg, one to three days prior to each treatment cycle. However, deviations from these schedules are often encountered during the course of the trial as patients may not be able to complete the questionnaire at the scheduled time. Consequently, analyses of HRQOL data frequently use 'completion time windows' around the expected completion time (31).

A certain number of days before and after the scheduled treatment cycle date may be allowed, so that all questionnaires completed within that period are assumed to belong to that particular cycle.

Such time intervals, rather than a single fixed visit, allow more flexibility in data collection, thereby minimizing missing data. The impact of these completion time windows was explored, involving RCTs of different cancer treatments, eg, radiotherapy and chemotherapy.

We examined whether the QLQ-C30 scores of cancer patients undergoing chemotherapy are affected by the specific time point before, during, or after treatment, at which the questionnaire is completed, and whether this could bias the overall treatment comparison analyses. Using linear mixed models for the analyses of longitudinal data (31), we found statistically significant differences ($P < .05$) for during and after treatment comparisons in these trials (17). For all three closed RCTs, the longitudinal mixed models resulted in a better fit when the 'completion time window' variable was included. However, differences were not sufficient to change treatment effects. While in this case accounting for time of completion did not alter the interpretation of treatment comparisons, findings might vary in other situations. Additional analyses are ongoing to replicate these results in radiotherapy clinical trials.

Discussion

The PROBE initiative used archived cancer clinical trials as the basis for pooled HRQOL data. Our exploratory work shows it is possible to pool trials from across the world in order to address key HRQOL issues. It demonstrated the statistically significant advantages that can be made when international and multidisciplinary scientists (clinicians, methodologists, and psychologists of multinational background and diverse scientific experience), such as those who comprised the PROBE Advisory Board, are brought together and join forces under a common scope. The PROBE members' specific mandate was to focus on the practical application of HRQOL in RCT research to critically explore important aspects of implementing HRQOL in cancer clinical trials. PROBE's contribution to clinical research was recognized when one PROBE publication was featured as a "notable advance" that successfully informed and changed clinical practice according to ASCO's annual report on progress against cancer, "Clinical Cancer Advances 2012" (20). We continue to encourage collaboration at every level and have raised these issues at the European Parliament, which has been very supportive of our initiative (32).

PROBE investigated critical topics with a major impact on future psychosocial care. Key outcomes were identified for improving the survival prognostication using HRQOL data, collected through the QLQ-C30. By exploring and identifying clusters of HRQOL problems, evidence of symptom interrelation was found that can benefit patients and lead to improvement of symptom management. The MID varied across brain and lung cancer patients, indicating there is need for additional empirical investigation of MIDs. Last, it provided valuable evidence on the importance of using completion time windows in the design of clinical trials and in the analysis of HRQOL outcomes. Yet, the variability of results and diversity of symptoms across cancer sites make the application and generalization of the findings challenging. PROBE will undertake research to estimate and compare MIDs on the QLQ-C30 across various patient populations that will provide guidelines for general interpretation.

However, we must not gloss over several methodological and statistical challenges and problems that had to be overcome during the course of pooling HRQOL and clinical data to conduct these retrospective analyses. The lack of availability of common indicators important for a full investigation of the research question was a major challenge (eg, disease stage, medical history). Variability across clinical trials on many important components, like the timing of assessments, made merging data a complex task. To address these challenges, we were forced to use complex methodological designs and techniques to account for heterogeneity (33). Dropout and nonadherence with HRQOL assessments at follow-up limited the available HRQOL data and presented considerable missing data challenges. The various data management systems used by different clinical trials organizations made pooling time consuming. Data privacy and ownership issues also made pooling difficult and, in some cases, precluded participation of interested researchers or groups. Equally challenging was the task of funding research dissemination via the most appropriate channels and media (ie, conference organization, press releases, etc.) to provide a greater understanding of HRQOL and anticancer treatment and boost public and clinical awareness. Infrastructure barriers and the lack of specialized staff in cancer-related HRQOL were other issues we had to overcome to ensure project sustainability. A limitation of our research relates to the disproportional representation of the study population because of the availability of large-scale studies within certain cancer types (eg, 3125 melanoma patients vs 78 esophageal cancer patients) (Figure 2). Another limitation is poor HRQOL compliance in cancer clinical trials and the lack of good longitudinal HRQOL data, which restricted our research mainly to baseline data. Going forward, one way to overcome the challenges that the pooled datasets present would be the use of standardized core clinical trial datasets for outcome measures, standardized collection and coding, and more clinical trial data sharing. Uniform publication guidelines (34–35) will allow HRQOL data to be more consistently presented.

Likewise, constraints caused by study population and HRQOL compliance difficulties are expected to be overcome in the future with the inclusion of more recent academic and industry RCTs with good compliance and longitudinal HRQOL data across various cancer sites.

In terms of future research, numerous analyses are being planned to make better use of the PROBE data and to help gain a better understanding of clinical trial data. Research topics under investigation include missing data, prognostic value of cancer-specific modules, MID in all cancer types, and joint modeling of longitudinal HRQOL (36) with clinical outcomes (such as overall survival, progression-free survival, and other biomarkers). Further research correlating HRQOL with survivorship and biomarkers (which have now become a standard component of EORTC clinical trials) will also be investigated when the necessary funds are secured. Long-term evaluations linking HRQOL to survivorship data are another goal.

The use of closed international clinical trials holds great promise for understanding patients' unmet psychosocial and HRQOL needs. This research program has shown that expensive prospective and lengthy studies are not always needed to answer specific research questions. Development of large-scale global collaborations such as the PROBE initiative has proven to be a valuable way of using existing data and benefits both patients and society by improving clinical care and understanding cancer. Such a repository of data will prove useful for many years to come. An automated way of adding new trial data to

enrich the HRQOL dataset will be developed by the EORTC with the support of our international collaborators. The growth of PROBE with 10 newly closed and fully published EORTC RCTs and the inclusion of pharmaceutical clinical trial data will provide an even larger and ever-growing database. This will allow us to answer more complex and overarching questions on key topics in oncology, such as institutional compliance. However, access to datasets and the merging of data are complex, with difficult processes; we need to be realistic about the challenges and expenditure. We invite other clinical trial researchers who have an interest in HRQOL research to work with us, join us, share data, and increase the pool of data, so many more important questions can be addressed and new questions can be developed.

In summary, we hope the PROBE initiative has shown that closed RCTs with HRQOL data can play an important role in the planning of future research, promote a better understanding of cancer care and the role of patient-reported HRQOL assessments (37), and extend our knowledge of methodological issues in HRQOL assessment. This initiative has demonstrated the benefits of international collaboration in the field of HRQOL. Cooperation contributes considerably to improving the efficiency and impact of research efforts, and only with increased collaboration across research groups can we address many of the outstanding HRQOL questions.

Funding

This work was supported by the European Organisation for Research and Treatment of Cancer (EORTC) Cancer Research Fund and an unrestricted academic grant from the Pfizer Foundation Global Health Partnerships Program, administered through the King Baudouin Foundation, for the EORTC Patient-Reported Outcomes and Behavioural Evidence (PROBE) project.

Notes

The funders had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

We thank the EORTC Headquarters, EORTC Clinical Groups, PROBE Advisory Board, NCIC-Clinical Trials Group, German Ovarian Cancer Study Group, all the Principal Investigators who helped us better understand the needs of cancer patients, which will ultimately lead to better patient care, all the patients who participated in these trials, Cheryl Whittaker, and Sheila Sanderson for editorial help.

References

- Bottomley A, Quinter C, Coens C, et al. Making better use of existing cancer data: Patient Reported Outcomes and Behavioural Evidence: a new international initiative. *Eur J Cancer Care*. 2009;18(2):105–107.
- 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.
- Luckett T, King MT, Butow PN, et al. Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. *Ann Oncol*. 2011;22(10):2179–2190.
- Efficace F, Osoba D, Gotay C, Sprangers M, Coens C, Bottomley A. Has the quality of health-related quality of life reporting in cancer clinical trials improved over time? Towards bridging the gap with clinical decision making. *Ann Oncol*. 2007;18(4):775–781.
- Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol*. 2008;26(8):1355–1363.

- Claassens L, van Meerbeeck J, Coens C, et al. Health-related quality of life in non-small-cell lung cancer: an update of a systematic review on methodologic issues in randomized controlled trials. *J Clin Oncol*. 2011;29(15):2104–2120.
- Zikos E, Ghislain I, Coens C, et al. Health-related quality of life in small-cell lung cancer: a systematic review on reporting of methods and clinical issues in randomised controlled trials. *Lancet Oncol*. 2014;15(2):e78–e89.
- Velikova G, Coens C, Efficace F, et al. Health-related quality of life in EORTC clinical trials — 30 years of progress from methodological developments to making a real impact on oncology practice. *Eur J Cancer Suppl*. 2012;10(1):141–149.
- Quinter C, Coens C, Mauer M, et al. Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. *Lancet Oncol*. 2009;10(9):865–871.
- Quinter C, Martinelli F, Vercauteren J, et al. Use of health-related quality of life and clinical data as prognostic tools for survival prediction in a subgroup of metastatic cancer patients. In: *46th Annual Meeting of the American Society of Clinical Oncology (ASCO)*, Chicago (USA), June 2010. *J Clin Oncol*. 2010;28:7s (suppl; abstr 9146).
- Quinter C, Martinelli F, Coens C, et al. A global analysis of multirial data investigating quality of life and symptoms as prognostic factors for survival in different tumor sites. *Cancer*. 2014;120(2):302–311.
- Ediebah DE, Coens C, Zikos E, et al. Does change in health-related quality of life score predict survival? Analysis of EORTC 08975 lung cancer trial. *Br J Cancer*. 2014;110(10):2427–2433.
- Quinter C, Maringwa J, Gotay CC, et al. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J Natl Cancer Inst*. 2011;103(24):1851–1858.
- Martinelli F, Quinter C, Maringwa JT, et al. Examining the relationships among health-related quality-of-life indicators in cancer patients participating in clinical trials: a pooled study of baseline EORTC QLQ-C30 data. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(5):587–599.
- Maringwa JT, Quinter C, King M, et al. Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Support Care Cancer*. 2011;19(11):1753–1760.
- Maringwa J, Quinter C, King M, et al. Minimal clinically meaningful differences for the EORTC QLQ-C30 and EORTC QLQ-BN20 scales in brain cancer patients. *Ann Oncol*. 2011;22(9):2107–2112.
- Ediebah DE, Coens C, Maringwa JT, et al. Effect of completion-time windows in the analysis of health-related quality of life outcomes in cancer patients. *Ann Oncol*. 2013;24(1):231–237.
- Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M. The EORTC QLQ-LC13: a modular supplement to the EORTC core quality of life questionnaire (QLQ-C30) for use in lung cancer clinical trials. *Eur J Cancer*. 1994;30A(5):635–642.
- Hahn EA, Cella D, Chassany O, Fairclough DL, Wong GY, Hays RD. Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clin Proc*. 2007;82(10):1244–1254.
- Roth BJ, Krilov L, Adams S, et al. Clinical cancer advances 2012: annual report on progress against cancer from the american society of clinical oncology. *J Clin Oncol*. 2013;31(1):131–161.
- 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.
- Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res*. 1993;2(3):221–226.
- Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. New Jersey: Lawrence Erlbaum Associates; 1988.
- 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.
- Bottomley A, Jones D, Claassens L. Patient-reported outcomes: assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicines Agency. *Eur J Cancer*. 2009;45(3):347–353.
- Giaccone G, Debruyne C, Felip E, et al. Phase III study of adjuvant vaccination with Bec2/bacille Calmette-Guerin in responding patients with limited-disease small-cell lung cancer (European Organisation for Research and Treatment of Cancer 08971-08971B; Silva Study). *J Clin Oncol*. 2005;23(28):6854–6864.
- Smit EF, van Meerbeeck JP, Lianes P, et al. Three-arm randomized study of two cisplatin-based regimens and paclitaxel plus gemcitabine in advanced non-small-cell lung cancer: a phase III trial of the European Organization for Research and Treatment of Cancer Lung Cancer Group—EORTC 08975. *J Clin Oncol*. 2003;21(21):3909–3917.
- Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol*. 2009;10(5):459–466.
- Van den Bent MJ, Carpentier AF, Brandes AA, et al. Adjuvant procarbazine, lomustine, and vincristine improves progression-free survival but not overall survival in newly diagnosed anaplastic oligodendrogliomas and oligoastrocytomas: a randomized European Organisation for Research and Treatment of Cancer phase III trial. *J Clin Oncol*. 2006;24(18):2715–2722.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102–109.

31. Fairclough DL. *Design and Analysis of Quality of Life Studies in the Clinical Trials*. London: Chapman and Hall/CRC; 2002.
32. Bottomley A, Kirby T, Bean J, Walker J. Moving the boundaries of international collaboration on clinical trials and QoL: experiences in oncology and legislation within the European Parliament. *Expert Rev Pharmacoecon Outcomes Res*. 2013;13(1):43–45.
33. Quinten C, Martinelli F, Coens C, et al. Investigating heterogeneity for survival in a pooled cancer cohort of more than 10,000 patients from 30 EORTC randomized trials. In: 17th Annual Conference of the International Society for Quality of Life Research. London, UK; Poster 150/1291, Oct 2010.
34. Calvert M, Blazeby J, Revicki D, Moher D, Brundage M. Reporting quality of life in clinical trials: a CONSORT extension. *Lancet*. 2011;378(9804):1684–1685.
35. Calvert M, Brundage M, Jacobsen PB, Schünemann HJ, Efficace F. The CONSORT Patient-Reported Outcome (PRO) extension: implications for clinical trials and practice. *Health Qual Life Outcomes*. 2013;11:184.
36. 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.
37. Santana MJ, Haverman L, Absolom K, et al. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. *Qual Life Res*. 2015;24(7):1707–1718.